

Development of the CRISPR/Cas9 Method for Use in *T. reesei*



Contents

1	Abbreviations.....	1
2	Introduction.....	2
2.1	Factors Affecting Cellulase Expression in <i>T. reesei</i>	3
2.2	CRISPR/Cas9	6
2.2.1	CRISPR/Cas9 Mechanism	7
2.3	The Time Before CRISPR/Cas9: HR, ZFNs, TALENs.....	9
2.4	CRISPR/Cas9 as a Research Tool.....	12
2.4.1	Non-transient Methods	12
2.4.2	Transient Methods.....	13
2.4.3	Catalytically Inactive Dead Cas9	15
2.4.4	gRNA Design	17
2.4.5	Off-target Effects	19
3	Project Aims.....	22
4	Materials and Methods.....	23
4.1	<i>T. reesei</i> Strains.....	23
4.2	Plasmids.....	23
4.3	Plates, Culture Mediums, and Cultivation Conditions.....	24
4.4	<i>T. reesei</i> Transformation	27
4.5	Target Gene Selection	29
4.6	Protein Extraction & Western Blot	29
4.7	Total RNA Extraction and qPCR	30
4.8	CRISPR/Cas9 Settings	31
4.8.1	Cas9 and sgRNA-ribozyme-complex Expression	31
4.8.2	gRNA Design and Production	35

4.8.3	dDNA Design.....	38
4.8.4	Protein Transformation.....	39
4.8.5	AMA1 Plasmid Transformation	40
4.9	Transporter Deletions	41
4.10	Enzyme Assays.....	41
4.11	Sequencing	42
5	Results.....	43
5.1	Plasmids.....	43
5.2	gRNA Design	44
5.3	gRNA Synthesis	45
5.4	<i>T. reesei</i> Continuous <i>Cas9</i> Expression & Ribozyme Constructs	45
5.5	AMA1 Transformation	53
5.6	<i>Cas9</i> Protein Transformation	55
5.7	Transporter Deletions	64
6	Discussion.....	70
6.1	CRISPR/ <i>Cas9</i> in <i>T. reesei</i>	70
6.2	Transporter Deletions	73
7	Conclusions.....	75
8	Acknowledgements	76
9	References.....	77

1 Abbreviations

AMA1 – Autonomous maintenance in *Aspergillus*

AmdS - Acetamidase

Cas9 - CRISPR-associated protein 9

CRISPR - Clustered regularly interspaced short palindromic repeats

CRISPRa – CRISPR activation

CRISPRi – CRISPR interference

crRNA – CRISPR RNA

dCas9 - Dead CRISPR associated protein 9

DSB – Double stranded break

DNA – Deoxyribonucleic acid

GFP – Green fluorescent protein

GOI – Gene of interest

gRNA – guideRNA

HR - Homologous recombination

IVT – *In vitro* transcription

ORF – Open reading frame

PAM – Protospacer adjacent motif

PCR – Polymerase chain reaction

Pyr4 – Orotidin-5'-phosphate decarboxylase

RNA – Ribonucleic acid

SSD – Single stranded break

sgRNA – Single guide RNA

tracrRNA – Trans-activating RNA

2 Introduction

Trichoderma reesei, an anamorph of *Hypocrea jecorina*, is a filamentous fungus widely used for producing industrial enzymes. The potential of *T. reesei* for industrial use lies not only in its ability to thrive under harsh industrial growth conditions, but also in its ability to produce vast amounts of enzymes, both endogenous and heterologous (Schuster, Schmoll 2010). The highest published level of secreted protein produced by *T. reesei* was over 100 g/L (Cherry, Fidantsef 2003). In addition to this, the secretion capacities of filamentous fungi are good and they are able to generate similar protein folding and modifications as other eukaryotes (Aro, Pakula et al. 2005). *T. reesei* is nowadays the most widely used filamentous fungus for heterologous protein production (Schuster, Schmoll 2010). Endogenously *T. reesei* produces vast amounts of cellulose degrading enzymes, cellulases and hemicellulases, which degrade plant biomass into glucose (Schuster, Schmoll 2010). This glucose is primarily used for cell growth by the fungi.

In 2014, 9 characterized cellulases, 15 characterized hemicellulases and at least 42 putative genes encoding (hemi)cellulolytic proteins had been identified in the genome of *T. reesei* (Häkkinen, Valkonen et al. 2014). The (hemi)cellulase production capacity of *T. reesei* is huge, and the regulatory pathways involved in their production have been studied extensively. A review article by Bischof et al. (2016) suggests that concentrating on the transcriptional regulation alone may not result in improved cellulase production and secretion in the current production strains, but pathways involved in the mRNA stability, translation, and enzyme secretion should be studied in more detail (Bischof, Ramoni et al. 2016).

While *T. reesei* has been used as a model organism for producing cellulases, the tools to genetically engineer this fungus are still slow and underdeveloped. A rapid way to accurately modify this organism's genome and gene expression levels could lead to faster development efforts in improving the production of cellulases and ultimately lead to reduced prices of the enzymes. Plant material rich in polysaccharides is routinely used to produce bioethanol for example, but the high cost of material pre-treatment and enzymes, such as cellulases, needed to turn the material into fermentable sugars brings up the price of bioethanol

(Himmel, Ding et al. 2007, Yang, Wyman 2008, Paloheimo, Haarmann et al. 2016). Rapid genome editing could help us better understand the process of plant material degradation and result in better enzyme yields or more effective enzymes lowering the costs of pre-treatment and production. As suggested by Martinez et al. (2008), *T. reesei*'s ability to degrade plant material doesn't seem to be a consequence of the production of multiple cellulases and hemicellulases. On the contrary, the full genome sequencing shows that *T. reesei* has only a few of these genes when compared to other fungi (Martinez, Berka et al. 2008). Discovering the underlying mechanisms that make *T. reesei* an extraordinary cellulase producer could mean great advancements in the cellulase production capacities.

To provide a tool to modify *T. reesei*'s genome accurately and quickly, CRISPR/Cas9 can be optimized for use with this organism. With the use of this tool, predicted regulatory mechanisms of *T. reesei* could be better understood. CRISPR/Cas9 not only provides the tools for accurate mutation introduction, but also the possibility of introducing multiple mutations at once. The system can also be used for adjusting gene expression levels. The main goal of this project was to find an effective way to introduce specific deletions into *T. reesei*'s genome with the CRISPR/Cas9 system and assess the effects of these deletions on the cellulase production efficiency.

2.1 Factors Affecting Cellulase Expression in *T. reesei*

Natural cellulosic breakdown into glucose by filamentous fungi is a key factor in the natural carbon cycle. Cellulose and hemicellulose are typically found in the plant cell wall together with pectin and lignin. Most organisms cannot degrade the plant cell walls and use (hemi)cellulose as an energy source, because the hydrolysis is difficult and involves multiple plant wall hydrolysing enzymes. *T. reesei* produces potentially dozens of (hemi)cellulolytic enzymes as well as pectinases and ligninases (Aro, Pakula et al. 2005, Häkkinen, Valkonen et al. 2014). This enzyme mixture gives *T. reesei* the ability to not only degrade the plant walls but to hydrolyse the (hemi)cellulose into glucose. This exact ability makes *T. reesei* the most popular biological platform for cellulase production.

Figure 1 summarizes three types of cellulases and their major functions in cellulosic degradation (Kumar, Singh et al. 2008). As *T. reesei* is widely used for cellulase expression, the regulation of the cellulase genes has been a hot topic of research. Inhibition caused by cellulase break-down products, including glucose and cellobiose, was already reported in the mid-1900's (Mandels, Reese 1960, Cohn, Horibata 1959). Glucose repression was identified as a type of transcriptional inhibitor already in 1997 (Aro, Pakula et al. 2005, Ilmén, Saloheimo et al. 1997).

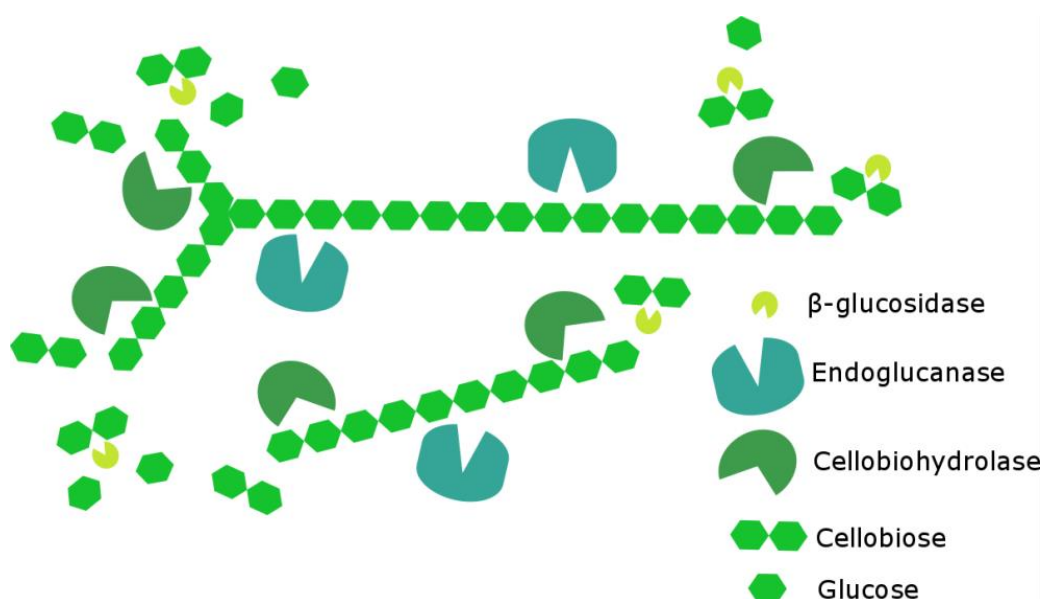


Figure 1 Cellulose degradation by cellulolytic enzymes. Endoglucanases randomly cleave the β -glycosidic bonds of the long cellulose chains. Cellobiohydrolases cleave cellobiose from the end of the cellulose chain and β -glucosidases cleave cellobiose into glucose monomers (Kumar, Singh et al. 2008).

Cellulase activities have been studied not only in regard to glucose but other carbon sources as well. In the QM6a *T. reesei* wildtype strain, D-lactose, Avicel and xylan induce the production and secretion of cellulases in liquid culture when 7 soluble and 7 insoluble carbon sources were tested (Dashtban, Buchkowski et al. 2011). The induction pattern changes and a wider range of carbon sources induce cellulase secretion in *T. reesei* mutant strains QM9414 and RUT-C30, which have been developed to express high amounts of cellulases (Dashtban, Buchkowski et al. 2011). However, these mutants have been created through classical mutagenesis and the exact mutations and genes responsible for the

increased cellulase production have not been fully understood (Seidl, Gamauf et al. 2008, Le Crom, Schackwitz et al. 2009, Vitikainen, Arvas et al. 2010). The studies on QM9414 and RUT-C30 strains suggest that the enhanced cellulase production capabilities are not due to single gene mutations, but involve multiple translocations, deletions, insertions, and single nucleotide variations (Le Crom, Schackwitz et al. 2009). From an industrial point of view the feedback inhibition caused by glucose may be the most significant disadvantage in using *T. reesei* as a cellulase producer. After all, *T. reesei* produces (hemi)cellulases for its own purposes: to hydrolyse cellulose and create glucose for growth. The cellulase production machinery itself is very efficient and the high production yields hinder cell growth. This is why *T. reesei* only produces cellulases under necessary conditions. Energy is not wasted on breaking down cellulose if a simple glucose carbon source is available.

The inhibition caused by extracellular glucose hinders industrial cellulosic breakdown as the production is halted once glucose is present in an adequate concentration to allow for efficient cell growth (Teugjas, Väljamäe 2013). Ilmén et al. (1997) have been able to diminish the glucose inhibition by mutating repressor binding sites in the genome (Ilmen, Onnela et al. 1996). RUT-C30 and QM9414 are more resistant to glucose inhibition than QM6A for example, and QM9414 derivatives are generally recognized as the best heterologous protein producers (Peterson, Nevalainen 2012, Paloheimo, Haarmann et al. 2016).

The cellulase regulatory system is a complicated network of both intra- and extracellular factors that are recognized by e.g. sugar sensors and sugar transporters. The regulation is highly dependent on the available carbon sources and the regulators involved vary accordingly. Some inhibitory effects caused by extracellular molecules have already been identified in mid-1900's (Cohn, Horibata 1959, Mandels, Reese 1960). For example, intracellular β -glucosidases seem to affect the expression level of a cellulase inducer *crt1* on lactose, but have no effect on *crt1* expression on cellobiose (Xu, Zhao et al. 2014).

The transcription factors involved in the cellulase expression pathway have been studied extensively, but the sensing systems of *T. reesei* remain poorly understood. There is very little understanding of how the fungus recognizes its environmental conditions. Some *T. reesei* sugar transporters or sensors have been identified, and their effects on the (hemi)cellulose production and secretion pathways have been studied in regards to some

carbon sources (Zhang, Kou et al. 2013, Delgado-Jarana, Moreno-Mateos et al. 2003, Ivanova, Baath et al. 2013).

Crt1 is one of the few known sugar transporters in *T. reesei* and has been shown to function as a lactose and sophorose transporter. It has been recognized as an important gene in the cellulase expression pathway, as it's suggested to induce *xyr1* (Zhang, Kou et al. 2013, Ivanova, Baath et al. 2013). *Xyr1* in turn can be referred to as the master regulator of cellulase and hemicellulase expression. It activates the cellulase production system and affects the expression levels of up to 2706 genes depending on the available carbon source (dos Santos Castro, de Paula et al. 2016). *Crt1* is suggested to act independently of intracellular sugar sensors, but affects *xyr1* activation on cellulose, sophorose, and lactose (Zhang, Kou et al. 2013).

On the other hand, *Saccharomyces cerevisiae* has functioned as a model organism for eukaryotic glucose sensing and pathways both induced and repressed by glucose have been published (Kim, Roy et al. 2013). However, *S. cerevisiae* is not a cellulolytic organism and the sensing systems and especially the response to different carbon sources are likely to be very different when compared to *T. reesei*. Many of the genes involved in the yeast glucose pathways are sugar transporters, sugar sensors and/or evolutionary derivatives of these. The type and number of transporters on the cell membrane is dependent on the amount of available glucose (Kim, Roy et al. 2013). In this sense, glucose itself functions as the regulator of its own uptake by the cells (Özcan, Johnston 1999). Conversely, only one glucose transporter, Trhxt1, has been described in *T. reesei* (Ramos, Chambergo et al. 2006). This suggests that there are other yet-to-be-identified glucose sensors and transporters present in *T. reesei*'s genome.

2.2 CRISPR/Cas9

CRISPR/Cas9 is a fairly new and powerful tool to be utilized in genome editing. The method made a rapid entrance to the field of editing eukaryotic genomes, including mammalian cells, yeast cells, and fungi after its discovery as a gene editing tool in 2012 (Pennisi 2013). CRISPR/Cas9 system makes use of a Cas9 protein and a guide RNA (gRNA). Cas9 binds the

gRNA, which carries a Cas9 associating sequence and a sequence complementary to the desired target. Cas9 is bound to the target site, defined by the gRNA's sequence complementarity according to the Watson-Crick base pairing. Cas9 will then introduce a double stranded break (DSB) to the binding site. By taking advantage of a cell's own DSB repair methods or by introducing a piece of DNA with homology around the point of the DSB, the genome can be altered with high accuracy of the mutation locus (Sander, Joung 2014).

In this section the type II CRISPR/Cas9 system of *Streptococcus pyogenes* is presented. The *S. pyogenes* CRISPR/Cas9 is the most commonly used and thus the best known (Qi, Larson et al. 2013). Many other CRISPR/Cas9 systems are known and some are emerging as alternative gene editing systems. For simplicity, this thesis focuses on the *S. pyogenes* system. This should be kept in mind as the mechanism and design aspects of other systems differ.

2.2.1 CRISPR/Cas9 Mechanism

The CRISPRs, clustered regularly interspaced short palindromic repeats, can be found in multiple different prokaryotic species. These repeat sequences are separated by a spacer sequence, which often matches the sequence of a bacteriophage. The first step of the CRISPR/Cas9 mechanism is adaptation. When a prokaryote is infected by a bacteriophage, a part of its sequence is inserted as a spacer to the CRISPR locus (Figure 2A) (Marraffini 2015). The CRISPR sequences are then transcribed into short RNAs each containing one spacer sequence also known as CRISPR RNAs (crRNA). A trans-activating RNA (tracrRNA) and the Cas9 protein are produced separately (Figure 2B) (Marraffini 2015). The tracrRNA is first associated with the crRNA creating a tracr:crRNA complex which in turn associates with the Cas9 protein (Figure 2C). In this complex the crRNA sequence will function as a guide and if an infecting agent's genome has complementarity to the crRNA sequence, Cas9 will be guided and bound to the infecting DNA, introduce a DSB and induce the degradation of the foreign DNA (Figure 2 D & E) (Sander, Joung 2014). The introduction of novel spacer sequences allows CRISPR/Cas9 to function as a form of adapting immune system in prokaryotes (Marraffini 2015).

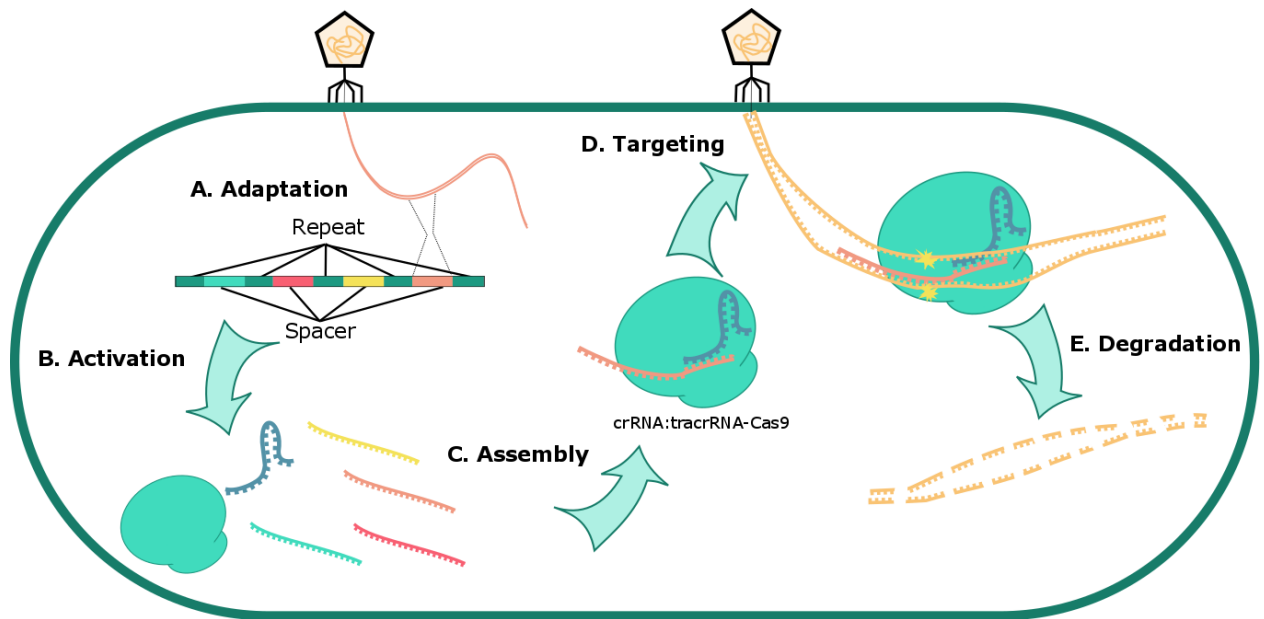


Figure 2 The CRISPR/Cas9 mechanism in the prokaryote *Streptococcus pyogenes* as described by Sander and Joung (2014) and Marraffini (2015).

The genome editing system of CRISPR/Cas9 makes use of a Cas9 protein and a crRNA:tracrRNA complex. A single guide RNA (sgRNA) is often used instead to simplify the system. The sgRNAs include both a Cas9 associating sequence and a specific target sequence in a single RNA molecule. For simplicity, the term guide RNA (gRNA) will be used to refer to either type of RNA.

Cas9 is only one type of CRISPR-associated proteins, but it is the most widely used one. The DSB inducing Cas9 is the only CRISPR-associated protein used in this study. Cas9 binds the gRNA's Cas9 associating sequence and is consequently bound to the target site, also known as the protospacer, defined by the gRNA's sequence complementarity and a protospacer adjacent motif (PAM). Cas9 will then introduce a DSB to the binding site. Figure 3 presents the Cas9 activity with an sgRNA. By taking advantage of a cell's own DSB repair methods or by introducing a piece of DNA with homology around the point of the DSB, the genome is altered at the CRISPR/Cas9 defined site (Sander, Joung 2014).

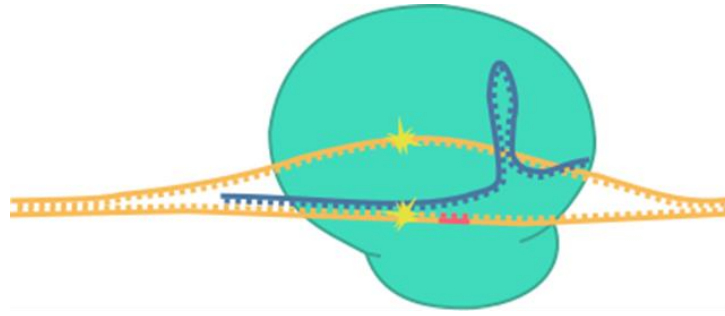


Figure 3 A schematic of Cas9 introducing a DSB to sgRNA and PAM specified target site. The red nucleotides represent the PAM-site.

2.3 The Time Before CRISPR/Cas9: HR, ZFNs, TALENs

The earliest method of introducing mutations to eukaryotic genomes was homologous recombination (HR). The early method was based solely on the homologous recombination mechanisms naturally present in the cells. In this method a piece of DNA including homologous flanks for the desired target is transformed into a cell, and the organisms own HR mechanisms will then integrate the DNA to the desired locus, defined by the homology. This resulted in random insertions and an overall low efficiency of the desired mutation. When using 14.3 kb homologous flanks a successful integration of the cassette frequency was 1 in 3×10^4 in embryonic stem cells (Capecchi 1989), but a more generally accepted recombination frequency is 1 in every 10^6 cells (Cathomen, Keith Joung 2008). Homologous recombination is still the basis of most genome editing methods, but on its own it's far too inefficient, unspecific and slow. In the case of creating genetically modified mice, it typically takes more than a year when using HR and the method doesn't work very well in human cells (Gupta, Musunuru 2014, Eiges, Schuldiner et al. 2001). Introducing genetic deletions, insertions and mutations into filamentous fungi is currently solely based on HR (Meyer 2008). The impairment of a fungi's NHEJ machinery radically increases the mutation efficiency. This makes HR a fairly efficient method in fungi (Krappmann 2007). The variance of the transformation method in use is in the DNA delivery method (Meyer 2008).

At the end of the 20th century a research group noted that introducing a DSB into the plant and mammalian genomes enhanced HR (Rouet, Smih et al. 1994, Puchta, Dujon et al. 1993). Less than ten years later, this lead to the emergence of zinc-finger nucleases (ZFNs) as tools to create DSBs and enhanced HR *in vivo* (Bibikova, Carroll et al. 2001). A Cys2-His2 zinc-finger

domain (ZF) is the most common DNA binding domain found in many organisms, including *H. sapiens*, *Drosophila*, *C. elegans*, and *S. cerevisiae* (Venter, Adams et al. 2001, Lander, Linton et al. 2001). Structural studies of zinc-fingers showed that target recognition is done by certain amino acids binding to DNA nucleotides of a single strand (Elrod-Erickson, Rould et al. 1996, Pavletich, Pabo 1991). Since then it was possible to design new zinc-finger domains with desired specificities. A review article by Beerli and Barbas (2002) indicates three different design types: rational design, combinatorial libraries, and a combination of the two (Beerli, Barbas 2002). When the ZF is fused with a catalytic endonuclease domain FokI, a targeted endonuclease ZFN is created. See figure 4A for a schematic description of ZFNs binding to their target. The first successful use of specifically target-designed ZFNs for genome editing resulted in 5,7% efficiency of disrupting the target gene when no donor DNA with homologous flanks was used (Bibikova, Golic et al. 2002). When the same set-up was done with a donor DNA an efficiency of up to 20% was reported (Bibikova, Beumer et al. 2003).

Another 10 years went by and transcription activator-like effectors nucleases (TALENs) were identified as potential genome editing tools (Christian, Cermak et al. 2010). Transcription activator-like effectors (TALEs) were first found from plant-pathogenic bacteria (Bonas, Stall et al. 1989). They bind to a specific DNA sequence defined by a repetitive amino acid sequence of the protein (Herbers, Conrads-Strauch et al. 1992). The variation of specificity is created by two varying amino acids in the repeat known as repeat variable diresidues (RVDs) or hypervariable aminoacids (Moscou, Bogdanove 2009, Boch, Scholze et al. 2009). The endonuclease activity is added to TALEs by fusing the same catalytic domain as in ZFNs, FokI, to the desired TALE (Christian, Cermak et al. 2010). This fusion protein is referred to as transcription activator-like effector nuclease (TALEN). See figure 4B for a schematic overview of TALENs binding to their target site. TALENs are reported to show indel efficiencies of up to 65% in flies (Wei, Liu et al. 2013).

The main difference in designing ZFNs and TALENs is that the DNA-binding domains recognize nucleotides in a different manner: A single ZFN unit binds three adjacent nucleotides whereas the TALEN RVDs recognize a single nucleotide (Mashimo 2014, Cermak, Doyle et al. 2011).

Both ZFs and TALEs can be used as transcription regulators by associating them with known activators or repressors, or simply by directing the protein in a site of the genome where it interferes with the transcription machinery assembly as described in two review articles (Beerli, Barbas 2002, Moore, Chandrabhas et al. 2014).

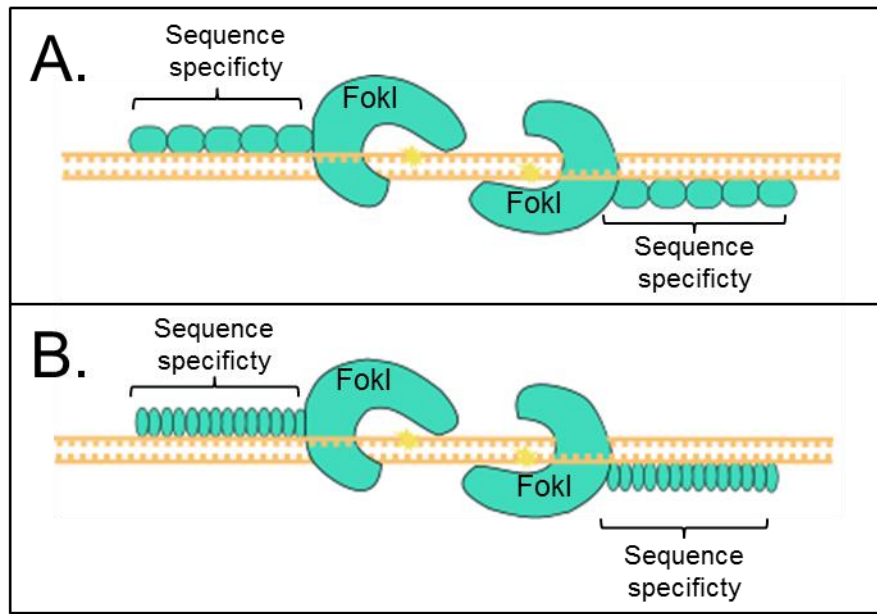


Figure 4. A schematic view of ZFNs (A) and TALENs (B). Both proteins are a fusion of a DNA-binding protein domain and a FokI endonuclease. The DNA binding domain is responsible for guiding the nuclease to the desired site of the sequence. Note that the FokI domain introduces a single stranded break. To obtain a DSB, two ZFNs or TALENs are needed to bind both DNA strands.

As recently as in 2013, a review article by Gaj et al. identified ZFNs and TALENs as “the forefront of genetic engineering” for their simplicity and flexibility (Gaj, Gersbach et al. 2013). However, a problem with both ZFNs and TALENs is the non-existing simple interaction between one amino acid and one nucleotide, as the adjacent aminoacids to the nucleotide-bound one seem to crucially alter the binding capacities (Dreier, Segal et al. 2000, Segal, Dreier et al. 1999). The situation is better with TALENs, as their binding depends on only two amino acids per nucleotide, but the binding capacities are not clear-cut (Morbiter, Römer et al. 2010, Cermak, Doyle et al. 2011). Also because of amino acids being the target-site recognizing subunits, using ZFNs and TALENs always requires designing and expressing a new protein for each target site. This is laborious, time-consuming, and thus rather expensive. Hence the main advantage provided by CRISPR/Cas9 is the fact that the catalytically active

protein does not need to be engineered for each target. Only the target-specific gRNA needs to be specifically designed for each site. Designing and producing a single gRNA molecule is easier, faster, and less expensive than doing the same for a protein. Modifications to the Cas9 protein itself allow for the use of CRISPR/Cas9 as a transcriptional regulator as well (see Section 2.4.3 Catalytically Inactive Dead Cas9).

2.4 CRISPR/Cas9 as a Research Tool

There are multiple ways to set up the CRISPR/Cas9 method. So far the scientific community has not found an organism in which the system would not be functional. The question mainly concerns how the components of the system should be delivered to each cell and organism type. These methods include expression vector transformations, RNA transformations, protein transformations, and a mixture of these. The preferred method depends on 1) the efficiency, 2) time-efficiency, 3) costs, and 4) the purpose of the study. Different types of methods are presented below, divided into non-transient and transient methods.

2.4.1 Non-transient Methods

In a non-transient method of CRISPR/Cas9 at least one of the components of the system is expressed in the organism of interest: *cas9* and/or gRNA. The non-transient method of expressing the CRISPR/Cas9 components *in vivo* is especially important when a catalytically inactive dCas9 is used (see Section 2.4.3 Catalytically Inactive Dead Cas9). The Cas9 protein can be expressed continuously or under an inducible promoter. The gRNA can be produced *in vitro* or *in vivo*. *In vitro* produced gRNA is simply transformed into a *cas9* expressing strain. Several kits for producing *in vitro* transcribed (IVT) gRNAs are on the market and custom-made gRNAs can also be ordered. As the review article by Bortesi et al. shows, *in vivo* production of gRNAs can be achieved with multiple different promoters, including type II and type III RNA polymerases for plant and mammalian cells (Bortesi, Fischer 2015). However, some promoters have specific sequence needs, e.g. the widely used U6 promoter needs a guanine (G) as the first transcribed nucleotide (Sander, Joung 2014). This problem can be

overcome by simply adding an extra G to the 5' end of one's guiding sequence (CRISPR Cas9 - gRNA Design. 2016). By using a ribozyme system presented by Gao et al., any promoter can be used for gRNA production and no restrictions to the transcribed sequence exist. In this system, the desired gRNA sequence is attached to ribozyme sequences in both 5' and 3' ends. After transcription, these ribozyme sequences will self-cleave (Gao, Zhao 2014). Figure 5 shows a schematic of the sgRNA-ribozyme-complex. According to Gao et al., this system can be used for both *in vivo* and *in vitro* production of gRNAs. This system has been shown to function in the fungi *Aspergillus niger* (Nødvig, Nielsen et al. 2015).

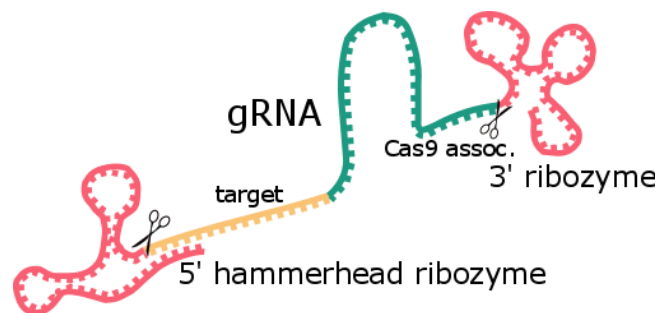


Figure 5 Visualization of sgRNA-ribozyme-complex post transcription from a chosen promoter.

2.4.2 Transient Methods

Using CRISPR/Cas9 as a transient mutation introducer is an attractive way to use the tool. An extensive study has shown that a continuously expressed *cas9* does not have detectable morphological or DNA damaging effects on mice on either tissue or cellular level (Platt, Chen et al. 2014). However, a slight decrease in cell viability has been reported when both *cas9* and gRNA were expressed in *S. cerevisiae* (DiCarlo, Norville et al. 2013) and potential toxic effects have been reported for *Chlamydomonas reinhardtii* (Jiang, Brueggeman et al. 2014). Some studies have also reported high off-target effects when Cas9 and gRNAs were produced *in vivo* (Ousterout, Kabadi et al. 2015, Fu, Foden et al. 2013) while a reduction in off-target events has been reported when using protein transfection rather than expression cassettes (Liang, Potter et al. 2015, Zuris, Thompson et al. 2015) (see Section 2.4.5 for off-

target effects). The idea of a continuously expressed endonuclease does awake concerns and the ease of introducing a mutation while leaving all other parts of the genome untouched is tempting. In a transient method the concentrations of both Cas9 and gRNAs can be finely adjusted, decreasing the likelihood of off-target effects (Hsu, Scott et al. 2013a, Pattanayak, Lin et al. 2013). Harmful off-target effects can also be decreased by using a Cas9 nickase instead of a DSB inducing Cas9 (see Section 2.4.3 Catalytically Inactive Dead Cas9). Another advantage of a transient method is that time-consuming cloning steps could be reduced and high-throughput gene editing made possible.

In the case of filamentous fungi, a major benefit to using a transient method is the possibility to accomplish gene deletions and insertions without a permanent marker. The lack of markers is a clear hindrance when modifying filamentous fungi as there are currently only 6 widely recognized markers and the recycling of these markers is very slow (Schuster, Schmoll 2010). Marker-free deletions were achieved in *P. chrysogenum* using a transient CRISPR/Cas9 method making this possible under certain conditions (Pohl, Kiel et al. 2016).

In a transient method, the question mainly remains whether it's possible to introduce a functional, yet only transiently present protein into the cells, as IVT gRNAs are already widely in use with the non-transient methods as well. There are a couple of ways in which a transient CRISPR/Cas9 system can be accomplished. One of these is a protein transformation, which is quite popular in mammalian cells. This method is based on the anionic nature of the Cas9-gRNA complex. Transfections of DNA, RNA, and protein into a mammalian cell can be done using a cationic liposomal reagent. The liposomal reagent fuses with the cell membrane and releases its contents into the cell. However, the electrostatic properties of proteins are unique, and the coupling of the cationic liposomal reagent and the coupling of any protein is not possible (Zuris, Thompson et al. 2015). In the case of Cas9, this problem can be solved by associating it with the gRNA prior to the transformation. The polyanionic gRNA gives the Cas9-gRNA complex a highly anionic nature, which allows for the coupling between the Cas9-gRNA complex and the liposomal reagent. A study by Zuris et al. found this protein transformation method to be highly efficient with up to 80% of the cultured cells being modified with few off-target events (Zuris, Thompson et al. 2015). Protein transformation has also been successful in the filamentous fungus *P. chrysogenum* although selection for successful transformants proved difficult (Pohl, Kiel et al. 2016). In the

study published by Pohl et al. (2016) the Cas9 protein and gRNA were first complexed together into a ribonucleoprotein and then transformed into a protoplasts without a liposomal agent (Pohl, Kiel et al. 2016).

In some filamentous fungi, an AMA1 plasmid can be used for transient expression of *cas9* and gRNA. AMA1 (Autonomously Maintained in *Aspergillus*) is a 6.1kb sequence, which allows for the independent replication of a plasmid (Mukherjee, Horwitz et al. 2013). A plasmid can be used to introduce an expression cassette either to be integrated to the genome or to be expressed from the plasmid. There are two benefits to using a plasmid with the AMA1 sequence: it's been shown to transfect *T. reesei* up to 60 times more efficiently than an integrative plasmid (Kubodera, Yamashita et al. 2002) and the plasmid is typically lost after some rounds of conidiation, thus allowing for transient expression (Mukherjee, Horwitz et al. 2013).

Protein expression can also be achieved through mRNA transformation. This has been done for *cas9* as well and is another method of transient expression (Wang, Yang et al. 2013). It holds similar benefits to AMA1 transformation: transformation is simple and the half-life of *cas9* mRNA is relatively short. No toxic effects have been noted when transforming IVT *cas9* mRNA and gRNA into mice zygotes (Wang, Yang et al. 2013). One benefit to using mRNA transformation rather than protein transformation may be the ease of the transformation. No liposomal reagent is needed and an mRNA is fast to produce with current IVT kits.

2.4.3 Catalytically Inactive Dead Cas9

The catalytic activity of Cas9 is a result of two endonuclease domains, HNH- and RuvC-like domains. Of these, the HNH domain cleaves the complementary strand of DNA and RuvC the noncomplementary strand (Jinek, Chylinski et al. 2012). By inactivating just one of these domains by introducing a point mutation, Cas9 can be used as a nickase, introducing only a single stranded break (SSB). This in some cases has been shown to provide more specificity and lower off-target effects for the DSB if two nicking Cas9s are targeted to a site by separate gRNAs. The method has shown great functionality in human cells (Cho, Kim et al. 2014, Shen, Zhang et al. 2014, Ran, Hsu et al. 2013). Using multiple gRNAs does however

increase the amount of off-targets, but this is generally not considered a problem with the nickases, as cells can efficiently repair an SSB (Ran, Hsu et al. 2013). If both of these domains are inactivated, Cas9 loses its catalytic activity and simply binds the DNA sequence specified by the gRNA. Qi et al. (2013) showed that the catalytically inactive Cas9, also known as dead Cas9 (dCas9) can be repurposed to repress bacterial gene expression by guiding it to the non-template strand of the gene of interest (GOI) open reading frame (ORF), the -35 box or the GOI or the promoter region (Qi, Larson et al. 2013). dCas9 can also be fused with a repressive domain, but using dCas9 alone is believed to interfere with the assembly of the transcriptional initiation machinery and thus is sufficient to repress a gene on its own (Qi, Larson et al. 2013). This transcriptional silencing was later shown to be functional in eukaryotic cells as well and is now known as CRISPR interference (CRISPRi) (Gilbert, Larson et al. 2013). Almost simultaneously with CRISPRi, CRISPR activation (CRISPRa) surfaced. In CRISPRa the dCas9 protein is fused with an activating domain. By guiding this upstream of the TATA-box or downstream of the transcriptional start site an increase in gene activity has been reported (Gilbert, Larson et al. 2013, Farzadfard, Perli et al. 2013, Konermann, Brigham et al. 2015). dCas9 has also been used to modify a cell's epigenetic marks by fusing it to a catalytic histone acetyltransferase or methyltransferase (Vojta, Dobrinić et al. 2016, Hilton, D'Ippolito et al. 2015). In addition, dCas9 fused to a green fluorescent protein (GFP) has been used to image and identify genomic repetitions (and non-repetitive sequences), telomere regions and movements, and gene loci on the genomic scale (Chen, Gilbert et al. 2013).

With dCas9 the guidelines for choosing a good gRNA target site are quite poor. Konermann et al. (2015) identified rough guidelines of choosing a CRISPR activation target site -200 - +1 bp from ORF, but otherwise the sites where CRISPRi and CRISPRa should be targeted have not been thoroughly studied and further studies could result in better repressional or activating effects (Konermann, Brigham et al. 2015). Gilbert et al. published a robust outline for target site design for dCas9, and found the CRISPRi and CRISPRa highly specific to their targets (Gilbert, Horlbeck et al. 2014). To the best of our knowledge, the use of dCas9 has not been reported in filamentous fungi.

These examples of the utilization of dCas9 exhibit the wide array of re-purposing the CRISPR/Cas9 technology beyond traditional gene editing at the sequence level. The magnitude of different modifications done with CRISPR/Cas9 since its discovery show that

it's had a huge impact. The surfacing of this technology has arguably revolutionized the field of gene editing in just four years.

2.4.4 gRNA Design

In this section, only the guidelines and programs using a Cas9 protein derived from *S. pyogenes* are presented. Guidelines may differ when another Cas protein is used. The design guidelines are also highly affected as new information about CRISPR/Cas9 off-target effects emerge. While the off-target predictions play a crucial role in the gRNA design, these are presented in more detail in Section 2.4.5, Off-target Effects.

gRNA design really is target site design. The sequence complementary to the gRNA is referred to as the protospacer sequence. The Cas9 associating part of the sgRNA sequence may differ to some extent in length and sequence, but remains mostly the same. Same applies to the tracrRNA if an RNA duplex is used instead of an sgRNA. However, studies about the Cas9 associating sequence of the gRNA are emerging, and slight modifications such as extending the RNA and removing repetitive thymines (Ts) can improve the CRISPR/Cas9 efficiency (Dang, Jia et al. 2015, Chen, Gilbert et al. 2013). It should however be noted that the improvements done to the RNA sequence are dependent on the gRNA production method. Removal of the repetitive thymines arguably improves the transcription efficiency, and would hence be unnecessary when using *in vitro* produced gRNAs (Dang, Jia et al. 2015).

The basic rules in choosing a target site are quite simple: any 20nt sequence followed by the PAM-site NGG is a target for CRISPR/Cas9 (5'-N₂₀NGG-3'). The *S. pyogenes* derived Cas9 can be targeted to the non-coding strand as well, where a target would be PAM-sequence CCN followed by any 20 nucleotides (5'-CCNN₂₀-3')(Doench 2016). The PAM-sequence is not included into the protospacer sequence, but is recognized by the Cas9 protein itself, making it essential that the target sequence is followed directly with the PAM-sequence (Doudna, Charpentier 2014).

Where the Cas9 is guided exactly is a broader question. When a knock-out by inducing a frameshift-mutation is desired, the target site should be in the beginning of a gene's ORF to effectively disrupt the reading frame. When using a dDNA with flanks, it's generally considered that the DSB should be close to where the homology starts, up to a 200 bp distance according to addgene.org (Cortez 2015). In reality the target-site uniqueness plays a crucial role in choosing the target. Depending on the application a unique site is often more important than the exact DSB locus.

Multiple programs for CRISPR target site design are available both in the web and as add-ons to genome editing programs. Naito et al. (2014) published a new web-based gRNA design tool and reviewed the previous ones as well (Naito, Hino et al. 2015). All programs function a little differently. For example in most programs the input data is the sequence of a targeted gene and as a result the program lists suitable target sequences to this gene and ranks them based on off-target effect calculations (e.g. Geneious CRISPR Finder). However, Cas-OFFinder by Bae et al. uses only the target sequence as the input data and simply calculates possible off-target effects for that particular target (Bae, Park et al. 2014). Ma et al. (2013) published a program that would take single nucleotide polymorphisms (SNPs) into account when finding suitable targets and predicts the secondary structure of a chosen target site in a full length sgRNA (Ma, Ye et al. 2013). Most programs let the user choose how strict off-target prediction measures to use. Li et al. suggest that a target site GC content should be $\leq 35\%$ and up to 3 bp indels between the guide and protospacer sequences should be avoided especially at 5' end, 3' ends or around 7–10 bp from PAM and more than 12 bp away from PAM (Lin, Cradick et al. 2014). To the best of our knowledge, none of the available programs lets the user specify the locus of the indels, hence removing all 3 bp indels, mismatches and restricting the GC content would likely result in very few possible target sites. In this sense all target sequences should be checked by hand as well.

A few very simple shortcomings are common to all gRNA design programs and gRNA design in general. It is still quite poorly understood how specifically the gRNA finds its target and which parts of the target sequence are the most crucial (see Section 2.4.5 for Off-target Effects). Another problem is with the integrity of the sequence databases. Especially when it comes to using the dCas9 as a regulator, the promoter and regulatory region annotations are crucial for successful gRNA design. Mohr et al. (2016) point out that databases and

annotations for the most commonly used research organisms are updated frequently and the sequence data in the gRNA design programs may not be up to date (Mohr, Hu et al. 2016). Furthermore, the difficulty of predicting precise target sites is not only a question of sequence and annotation quality but also of individual differences. SNPs and random mutations in an individual's genome may result in unexpected and, in practice, impossibly predictable off-targets (Yang, Grishin et al. 2014).

Currently, gRNA design is done with the best of our knowledge, but further studies and more efficient sequencing and annotating technologies are needed before, if ever, any program can design perfect gRNAs. The lack of dependable sequence and annotation information is especially a problem when one targets the genome of non-model-organism or human genomes (Mohr, Hu et al. 2016). Currently the best approach is to test multiple different gRNAs to find a good one.

2.4.5 Off-target Effects

In all genome editing methods the specificity of a mutation is crucial. In a perfect method, the mutation is created directly at the desired site and only at the desired site. Ever since CRISPR/Cas9 made an entrance to the genome editing field, there's been a question of its specificity. The notion that CRISPR/Cas9 effectively creates a DSB 3 bp upstream of a PAM-site guided by the gRNA sequence was already shown early on (Jinek, Chylinski et al. 2012). In the case of CRISPR/Cas9 the question of concern really is if the DSB is introduced at unwanted sites in the genome.

This section is closely related to Section 2.4.4, gRNA Design, as all knowledge of CRISPR/Cas9 off-target effects should be taken into account in proper gRNA design.

Fu et al. (2013) published a study that showed that up to five mismatches in the gRNA target sequence are tolerated (Fu, Foden et al. 2013). They also showed that the specificity is not simple to predict. In their study, four different genes in the human genome were targeted and computationally identified potential off-target sites were studied for indel mutations. They found no off-target mutations in predicted sites for two of the genes (*Rnf2* and *Fancf*),

and up to 12 for the other two (*Vegfa* and *Emx1*). Another study by Cradick et al. (2013) found significant off-target effects with only a 10-12 bp match upstream of the PAM sequence to the gRNA sequence and one target that was cleaved when there were mismatches in the two bases proximal to the PAM site (Cradick, Fine et al. 2013). Cho et al. (2014) also showed off-target effects in predicted sites when a mismatch was introduced in the 6 nt sequence proximal to the PAM-site (Cho, Kim et al. 2014). This so-called seed sequence of the guide, defined as the 6-12 nt sequence directly upstream of the PAM-site was previously thought to be very specific to the target site. In addition to mismatches, the protospacer sequence seems to tolerate up to 4 bp long deletions and insertions by creating a bulge in the RNA or the target sequence (Lin, Cradick et al. 2014).

More high-throughput analysis of off-target effects has been done using a combination of ChIP-seq and computational prediction of off-target sites. O'Geen et al. (2015) analysed 1200 genomic loci and found only one off-target mutation in addition to the desired target mutation (O'Geen, Henry et al. 2015). Similar results were published by Wu et al. (2014), where out of 295 dCas9-bound sequences only one site other than the intended one was mutated by Cas9 (Wu, Scott et al. 2014). Both groups suggest that Cas9 binding and cleavage are separate events, and that the cleavage only takes place when binding is strong i.e. the protospacer matches the gRNA sequence well. Cencic et al. (2014) published similar results, but with only 43 ChIP identified target sites (Cencic, Miura et al. 2014). The results obtained using ChIP-seq analysis are very promising for the case of genome editing especially for the transient CRISPR/Cas9 methods, but raises concerns for continuously expressed *cas9* systems and the use of dCas9. Even when no cleavage takes place, both Cas9 and dCas9 can alter the expression levels of sites where they're bound (Duan, Lu et al. 2014). Wu et al. (2014) also pointed out that 70% of the off-target sites identified in their analysis as Cas9 binding sites were active genomic sites (Wu, Scott et al. 2014). What should be kept in mind though is that all of these studies reported a notably higher binding affinity to the target site than to any off-target site (Wu, Scott et al. 2014, Cencic, Miura et al. 2014, O'Geen, Henry et al. 2015). This would suggest for fairly transient attachment and detachment to and from the DNA, which may not affect the expression levels notably. Additionally, Gilbert et al. (2014) suggest that when using the dCas9 for CRISPRi or CRISPRa, the method is very specific

because the site where dCas9 must be bound to affect transcription is quite precise (Gilbert, Horlbeck et al. 2014).

Multiple studies on CRISPR/Cas9 off-target effects have identified the importance of the NGG-PAM for Cas9 and that different PAM sites are well discriminated (Cradick, Fine et al. 2013, Cho, Kim et al. 2014, Hsu, Scott et al. 2013b, Sternberg, Redding et al. 2014). No cuts were detected in completely complementary gRNA-target sequences, also *in vitro*, if another PAM site was present or a PAM-site was lacking. Furthermore, Sternberg et al. (2014) published a predicted method of how Cas9 finds its target (Sternberg, Redding et al. 2014). In this prediction the initial collision of Cas9 and a PAM-site initiates the RNA-DNA complementary alignment. They suggest that because this alignment begins directly upstream from the PAM-site, mismatches are less tolerated proximal to the PAM-site as the alignment is terminated and Cas9 dissociated more likely due to early mismatches (Sternberg, Redding et al. 2014).

Cho et al. (2014) identified a clear reduction in off-target effects when using an *in vitro* produced tracr:crRNA complex and using two additional guanines in the guide 5' sequence (5'-GGN₂₀-3') instead of *in vivo* produced sgRNAs with a 5' proximal guanine as part of the guiding sequence itself (Cho, Kim et al. 2014). Fu et al. published a study showing that a shorter protospacer sequence creates less off-target mutation than a 20 nt long (Fu, Sander et al. 2014). They showed that a 17-18 bp long protospacer sequence (truncated from 5' end, 5'-N₁₇₋₁₈NGG-3') in the gRNAs induced very low or undetectable off-target effects with just a 1-2 bp mismatch in the target sequence.

Although many questions about the off-target effects remain unanswered, it's clear that these should be taken into account when using CRISPR/Cas9 and analysing modified cell lines. As more studies emerge, we gain better understanding of choosing more and more specific guide RNAs and can use better methods to deliver the CRISPR/Cas9 induced modifications.

3 Project Aims

Three putative and one known sugar transporter or sensor genes were deleted and their effect on *T. reesei*'s ability to produce and secrete cellulases were assessed in an effort to find new regulatory factors that could be useful in a commercial fungal enzyme production platform.

As the CRISPR/Cas9 system requires several components, it may not be a faster method for performing individual gene mutations in *T. reesei*. As mutations are usually introduced by naturally occurring homologous recombination, the specificity aspect may not provide a significant benefit either. The full potential of CRISPR/Cas9 lies in its ability to introduce multiple different mutations at once. As *T. reesei* is a fairly slow organism to work with, this could provide a significant benefit when compared to the traditional genome editing methods. Thus, the goal of this project was to find the fastest and simplest way to use CRISPR/Cas9 in *T. reesei*, as well as, test the feasibility of performing multiple gene mutations at once with CRISPR/Cas9 in *T. reesei*.

This project began with the deletion of *T. reesei*'s sugar transporter gene *crt1* and a putative sugar transporter *cp129*. The deletions were attempted by producing all components of the CRISPR/Cas9 *in vitro* and transforming them into *T. reesei* (Cas9, gRNA and donor DNA). Another approach used was to test *in vivo* production of either Cas9 or gRNA or both.

Once a successful method of CRISPR/Cas9 in *T. reesei* was established, the deletions were done simultaneously to see if introducing multiple gene editions at once was feasible. Currently there are no reports of introducing multiple deletions in *T. reesei* at once with HR. Successful introduction of a double deletion can save approximately 3 months of the time used in strain generation.

4 Materials and Methods

4.1 *T. reesei* Strains

The *T. reesei* strains used in this study are listed in Table 1. M127 is a *pyr4*⁻ version of M124 and M123 is a *pyr4*⁻ version of M44. M717 is an mCherry indicator strain for cellulase activity.

Table 1 *T. reesei* strains used in this study.

Strain	Details
M123	<i>Mus53</i> ⁺ , <i>pyr4</i> ⁻
M124	<i>Mus53</i> ⁻ , <i>pyr4</i> ⁺
M127	<i>Mus53</i> ⁻ , <i>pyr4</i> ⁻
M717	pCbh1- <i>mCherry</i> , <i>pyr4</i> ⁺
M1692	pCbh1- <i>mCherry</i> , <i>pyr4</i> ⁻

4.2 Plasmids

Table 2 collects the plasmids used in this study. *T. reesei* optimized *cas9* expression (TROCas9) vector was a gift from the Zou Laboratory (Liu, Chen et al. 2015) and pCPL23-pCPL29 were constructed by C. Landowski at VTT.

Table 2 Plasmids used in this study.

Plasmid name	Contents	Bacterial selection	<i>T. reesei</i> selection
pCPL23	<i>Y. lipolytica</i> optimized <i>cas9</i> expression vector	Ampicillin	Hygromycin
pCPL25	<i>Y. lipolytica</i> optimized <i>cas9</i> expression vector	Ampicillin	Acetamide
pCPL26	<i>Y. lipolytica</i> optimized <i>cas9</i> expression vector	Ampicillin	Pyr4
pCPL27	Sugar transporter 1 deletion vector	Ampicillin	Pyr4
pCPL28	Sugar transporter 2 * deletion vector	Ampicillin	Pyr4
pCPL29	Sugar transporter 3 deletion vector	Ampicillin	Pyr4
TROCas9	<i>T. reesei</i> optimized <i>cas9</i> expression vector	Ampicillin	Hygromycin
pRS426	<i>S. cerevisiae</i> cloning vector with URA3 selection	Ampicillin	various

*Published as *stp1* (Zhang, Kou et al. 2013)

4.3 Plates, Culture Mediums, and Cultivation Conditions

Tables 3 and 4 describe the liquid media and agar plates used in this study. *T. reesei* was cultivated in two different media, SGE-lactose and glucose, to collect culture supernatant for enzymatic assays. Cultivation volume and time of incubation in +28 °C is specified for each assay. For DNA, RNA, and protein extraction from liquid culture SGE-lactose was used. Transformation plates and top agar were used for *T. reesei* transformations, and selection agent was chosen according to the transformed components. After 9-14 days in +28 °C (14-21 days for amds selection), transformed clones were streaked on TrMM-triton plates

with appropriate selection and incubated in +28 °C for 5-8 days. Successful transformants would be sporulated on a potato dextrose (PD-)plate for 4-7 days, a 2 ml spore stock collected through cotton-tip filtration in 20% glycerol, 0,8% NaCl and 0,025 Tween 20 and plated as single spores on TrMM-triton plates. Single spore clones would again be streaked on TrMM-triton plates for screening. Each strain was cleaned twice through a single-spore stage on TrMM-triton.

E. coli was cultivated in 5 ml LB medium supplemented with 100 µg/ml ampicillin for plasmid propagation. In the case of TOPO cloned gBlocks, 50 µg/ml kanamycin was used. Plasmids were collected with MiniPrep kit (ThermoFisher, USA) after an overnight incubation in +37 °C.

YPD media was used to cultivate *S. cerevisiae* ATCC 90845 overnight for transformation. SCD-URA plates were used for *S. cerevisiae* transformation selection. Successful transformants were collected from SCD-URA plates 3 days after incubation in +30 °C.

Table 3 Liquid culture mediums used in this study

Media Name	Details
SGE-lactose	2% SGE, 4% lactose with di ammonium citrate, 100 mM PIPPS, pH 4.8
Glucose	3% glucose, 1% yeast extract, pH 4.8
YPD	20 g/l Bacto peptone, 10 g/l yeast extract, 20 g/l glucose
LB	Luria Broth: Tryptone (10g/l), yeast extract (5g/l), NaCl (0,5g/l), pH 7.5.

Table 4 Agar plates used in this study

Plate name	Contents
Pyr4 transformation plates	KH ₂ PO ₄ (15 g/L), (NH ₄) ₂ SO ₄ (5 g/L), glucose (20 g/L), trace elements* (1 ml/L), sorbitol (182,2 g/L), granulated agar (18 g/L), MgSO ₄ (2,4 mM), CaCl ₂ (4,1 mM), pH 5.5.
Hygromycin transformation plates	Same as pyr4 transformation plates but with hygromycin (125 µg/ml).
Hygromycin-amds transformation plates	Same as pyr4 transformation plates, but with hygromycin (125 µg/ml), 10 mM acetamide, 12,6 mM CsCl.
TrMM-triton	Same as pyr4 transformation plates, but without sorbitol and with Triton X-100 concentration 0,1%.
TrMM-triton-hygro and TrMM-triton-amds	Same as TrMM-triton plates, but with hygromycin (125 µg/ml) or 10mM acetamide and 12,6mM CsCl..
Top agar	Same as pyr4 transformation plates, but 3% agar, supplemented with hygromycin (125 µg/ml) and/or CsCl (12,6 mM) according to selection.
PD-plate	potato starch (4g/l), dextrose (20g/l), granulated agar (18g/l), pH 5.5
LB-plate	Tryptone (10g/l), yeast extract (5g/l), NaCl (0,5g/l), BiTek agar (20g/l), pH 7.5. Supplemented with ampicillin (100µg/ml) or kanamycin (5µg/ml).
SCD-URA	2% galactose, BiTek agar (20 g/l), Difo yeast nitrogen base without amino acids (0,64 g/l), 20 mg/l R, H, M, 30 mg/l I, L, K, T, 50 mg/l F, 100 mg/l E, 150 mg/l V, 400 mg/l S, 40 mg/l Adenine.

*FeSO₄ x 7 H₂O (5 g/L), MnSO₄ x H₂O (1,6 g/L), ZnSO₄ x 7 H₂O (1,4 g/L), CoCl₂ x 6 H₂O (3,7 g/L)

4.4 *T. reesei* Transformation

The *T. reesei* transformations were done as described by the Penttilä et al. (1987) publication (Penttilä, Nevalainen et al. 1987). In this method a cassette for homologous recombination in *T. reesei* was assembled using *Saccharomyces cerevisiae* ATCC 90845. PCR amplified fragments including the gene of interest (GOI), fungal marker cassette and homologous flanks to the desired genomic site of *T. reesei* were transformed into ATCC 90845 with a linearized pRS426 backbone. *S. cerevisiae* transformants were plated on SCD-URA plates and colonies were collected in 0,9% NaCl after three days of incubation in +30 °C. DNA was rescued through cell lysis with plasmid release buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl (pH 8), 1 mM Na₂ EDTA) and glass beads followed by phenol-chloroform extraction. Collected DNA was then transformed into *E. coli* TOP10 cells from which plasmids were extracted with MiniPrep kit (ThermoFisher, USA). Successfully assembled plasmids were screened first with restriction analysis and then sequencing. Plasmids were amplified in *E. coli* TOP10. Figure 6 shows a schematic of an *E. coli* extracted plasmid.

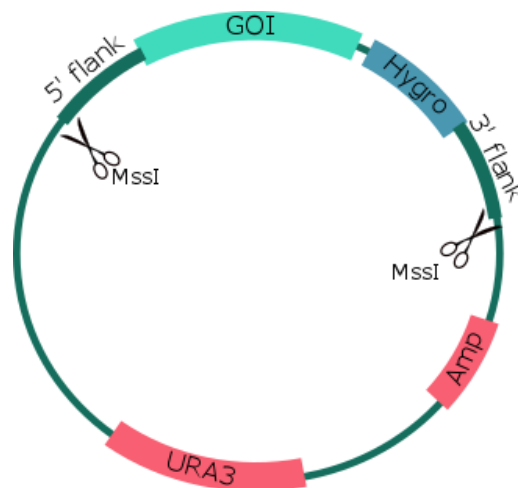


Figure 6 A schematic of the vector used for assembling and amplifying DNA for *T. reesei* transformation. In the case of a deletion vector, only a fungal marker cassette is located between flanking regions. Markers used in this study are hygromycin, pyr4 and acetamide.

10 µg of this DNA was digested with MssI to form a linear product (Figure 7A). The digested DNA was then treated with T7 exonuclease to create 5' overhangs to enhance the transformation efficiency (Figure 7B). This treatment was done to all traditionally

transformed DNA as well as marker dDNA used in the CRISPR/Cas9 transformations. No further purification or buffer removal steps were taken prior to transformation and the whole 10 µg of DNA was used for *T. reesei* transformation. In the case of double deletions, the two restricted and T7 exonuclease treated deletion cassette fragments were mixed together and half of the volume was used for transformation immediately. Hence the total DNA amount remained the same as in single fragment transformation.



Figure 7 T7 DNA exonuclease treatment prior to transformation. **A.** Cassette cut out of plasmid backbone with MssI restriction enzyme including the GOI, marker, and homologous flanks to the site of interest in *T. reesei* genome. **B.** T7 treated cassette with single stranded 5' overhangs.

Transformation was done through PEG 6000 mediated protoplast transformation. 10 µg of the pre-treated DNA was incubated with 250 µl $1-5 \times 10^7$ protoplasts/ml for 20 min on ice. 2 ml of 25% PEG 6000 – 50 mM CaCl_2 – 10 mM Tris-HCl pH 7.5 was mixed with the protoplast-DNA mixture to initiate DNA intake. Mixture was incubated for 5 min in room temperature. 4 ml 1.2M sorbitol – 10 mM CaCl_2 – 10 mM Tris-HCl pH 7.5 was added and mixed by inversion. The transformants were then plated on transformation plates with top agar, both with appropriate selection agents. After 9-14 days of incubation in +28 °C colonies were streaked on triton plates with appropriate selection. 5-7 days later DNA could be collected and screening for correct clones was done using Phire Plant Direct PCR kit (ThermoFisher, USA). PCR screening was done for the GOI open reading frame (ORF), insertion site ORF, and 5' and 3' integrations. For deletions, no GOI ORF screening could be done.

4.5 Target Gene Selection

In this study four gene deletions are introduced into the *T. reesei* genome. *Crt1* is briefly described in Section 2.1 Factors Affecting Cellulase Expression in *T. reesei*. *Crt1* is likely a carbon source transporter which, under certain conditions, regulates the so-called cellulase expression master regulator *xyl1*. This function makes *crt1* an interesting part of the cellulase expression system.

In addition to *crt1*, three putative *T. reesei* sugar transporters are deleted. These genes will be referred to with their deletion cassette plasmid names *cpl27*, *cpl28*, and *cpl29*. *Cpl28* (also known as *stp1*) has been reported to play a role in cellobiose and glucose transporting (Zhang, Kou et al. 2013). *Cpl27* and *cpl29* were found to have sugar transporter like homology according to the Joint Genome Institute (JGI) and were thus chosen as interesting candidates for understanding carbon source sensing and transportation through the cell membrane.

Crt1 and *cpl29* were deleted with the traditional method as well as the CRISPR/Cas9 method. Due to the time it took to set up the correct conditions for CRISPR/Cas9 in *T. reesei*, *cpl27* and *cpl28* were deleted through the traditional method only.

4.6 Protein Extraction & Western Blot

T. reesei intracellular proteins were extracted from 4-5 day old culture mycelium in order to determine Cas9 protein expression. Mycelia were collected from two 3 ml cultures in 24-well plates. Mycelia were collected by filtration through 1 micron glass microfiber Whatman filters (GE Healthcare, UK) and snap frozen in liquid nitrogen. Mortar and pestle were used to grind the mycelia into fine powder. 5 mM sodium-phosphate buffer, supplemented with EDTA-free protease inhibitor, was mixed with the mycelia powder. The sample was centrifuged 25 min at 14000 rpm, +4 °C and 10 min 5000 rpm, +4 °C. Supernatant was used as a protein sample for western blot. Some initial samples were prepared by mixing together mycelia and loading dye followed by 5 min incubation at +95 °C. The preparation style of the samples is indicated in the results.

For Cas9 protein detection, 15 µl of protein sample and 5 µl 4 × loading dye were incubated at +95 °C for 5 min and ran into a precast CriterionTMTGXTM 4-20% SDS-PAGE gel (Bio-Rad, USA) 250 V, 35 min. Proteins were transferred onto a Trans-Blot[®]TurboTM Transfer Pack 0,2 µM nitrocellulose film (Bio-Rad, USA) by Trans-Blot[®]TurboTM Transfer system's Midi program, 2,5 A, 25 V, 7 min (Bio-Rad, USA). The membrane was incubated with a 1:10 000 diluted rabbit anti-CRISPR-Cas9 antibody (ab204448, Abcam, UK) for 45-60 min. After washing the primary antibody, a 1:30 000 diluted Goat anti-rabbit IRDye 680RD secondary antibody (925-68071, Li-Cor Biosciences, USA) was added for at least 1 hour. After at least 30 min of wash with TBST and a rinse with TBS the membrane was imaged with Odyssey CLx 700 nm channel (Li-Cor Biosciences, USA).

4.7 Total RNA Extraction and qPCR

T. reesei total RNA was extracted from 4-day old 3 ml 24-well-plate cultures. Mycelia was collected by filtrating 1,5 ml of the culture through 1 micron glass microfiber Whatman filters (GE Healthcare, UK) and immediately frozen in liquid nitrogen. RNA extraction was done with RNeasy Plant Mini Kit (Qiagen, Netherlands). For storage, the RNA was frozen in liquid nitrogen and placed in -70 °C freezer.

cDNA was transcribed from total RNA either according to Transcriptor First Strand cDNA Syntesis kit (Roche, Switzerland) with dT₁₈ primers or according to QuantiTect Reverse Transcriptase kit (Qiagen, Netherlands) using the kit's primer mix which included a mixture of universal dT₁₈ primers and random primers. For sgRNA analysis primer PP145 was used to ensure transcription as the sgRNAs do not contain a polyA tail and may be too small for detection by random primers.

qPCR was done with Roche Light Cyclers SYBR Green I Master kit. All samples were diluted 1:10, 1:20 and 1:50 for qPCR analysis. Analysis was done with LightCycler[®] 480 II Standard SYBR Green I protocol (Roche, Switzerland). Table 5 summarizes primers used for detection.

Table 5 qPCR primers used to detect sgRNA, *cas9*, and *gpd1* transcripts.

Primer name	Detects
PP144 ¹ -PP145 ²	any sgRNA
PP181 ³ -PP182 ⁴	<i>Y. lipolytica</i> optimized <i>cas9</i>
C144 ⁵ -C145 ⁶	<i>T. reesei</i> optimized <i>cas9</i>
Gpd1Fw ⁷ -Gpd1Rv ⁸	<i>Gpd1</i> (housekeeping gene)

¹ GTTTTAGAGCTAGAAATA

² AGCACCGACTCGGTGC

³ GCTCTACCTCTACTACCTCCA

⁴ GTTCTTCATCTTCTTAACCACCTC

⁵ TCAACACCGAGATTACGAAG

⁶ CTTAATGATCTTCAGGAGGTC

⁷ TCCATTCTGTGTCCTACC

⁸ AGATACCAGCCTCAATGTC

4.8 CRISPR/Cas9 Settings

The sections below present the different settings used to set up the CRISPR/Cas9 method for *T. reesei*, including construction of continuous expression systems, dDNA design, gRNA design and production, and protein transformation.

4.8.1 *Cas9* and sgRNA-ribozyme-complex Expression

To introduce both gRNA and *cas9* expression in a single strain, a base vector pCPL34 was designed. Figure 8 shows a schematic design of this vector.

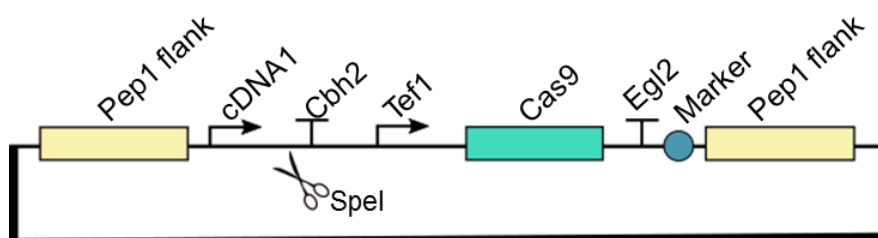


Figure 8 Design of base vector pCPL34. *Y. lipolytica* codon optimized *cas9* is expressed under a *tef1* promoter and terminated by *egl2*. A *SpeI* restriction site is located between the *cdna1* promoter and *cbh2* terminator, upstream of the *cas9* expression cassette, to allow for simple insertion of a desired sgRNA. Homologous recombination is directed to *pep1* locus by ~1000 bp flanking regions. Hygromycin marker was used.

Cas9 with its promoter and terminator, the *cdna1* promoter and *cbh2* terminator as well as the *pep1* flanks and hygromycin marker were introduced to the pRS426 plasmid by using PCR primers with appropriate 30 bp flanking regions to the plasmid. PCR fragments were recombined with the plasmid backbone in *S. cerevisiae* ATCC 90845. The plasmid was assembled in two sequential steps: first the *cas9* along with its promoter, terminator, selection marker and *pep1* flanks were amplified by PCR from pCPL23 with appropriate flanks and assembled into pRS426. An *Ascl* restriction site was included upstream of *tef1* promoter. After identifying a correct plasmid the *cdna1* promoter and *cbh2* terminator, separated by a *SpeI* restriction site, were added with the same method of using flanking PCR primers to *pep1* 5' flank and *tef1* promoter sequences and recombined to the *cas9* expression cassette linearized with *Ascl*. Screening for successful recombination was done with restriction analysis and sequencing. The *SpeI* restriction site allowed for the insertion of any desired sgRNA-ribozyme-complex downstream of the *cdna1* promoter. By restricting the plasmid with *SpeI* an sgRNA-ribozyme-complex could be introduced to the base vector by yeast-mediated homologous recombination. Figure 9 shows a schematic of the plasmid with an sgRNA-ribozyme-complex expression included. The different sgRNA-ribozyme-complexes were ordered as ready-made gBlocks (Integrated DNA Technologies, USA) with appropriate 30 bp flank regions to the *cdna1* and *cbh1* sequences. The *SpeI* site was abolished with the gBlock introduction.

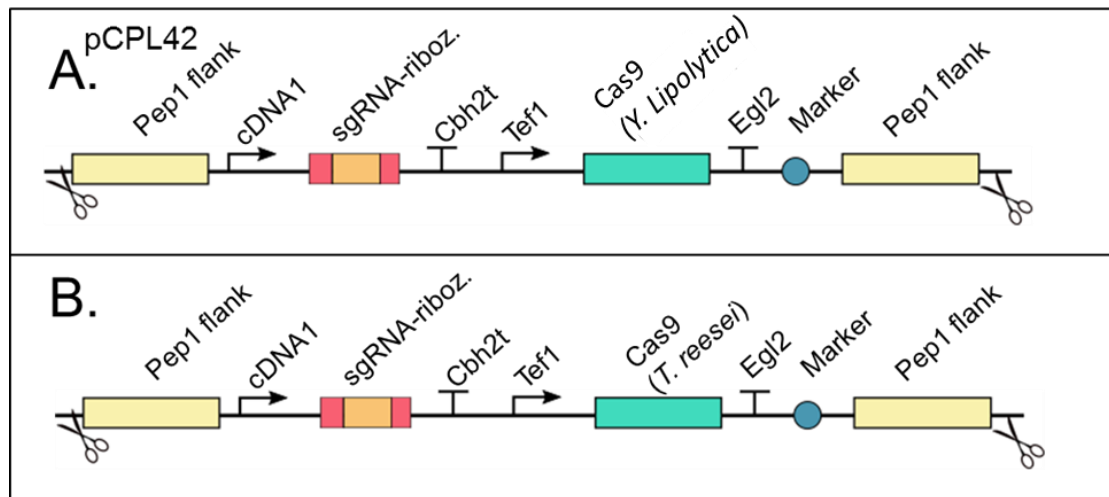


Figure 9 Base vector with an sgRNA expression cassette included. **A.** pCPL34 with sgRNA-ribozyme expression upstream of *tef1* promoter. sgRNA-ribozyme expressed under *cdna1* promoter, terminated by *cbh2t* terminator. **B.** pCPL49 with sgRNA-ribozyme expression upstream of *tef1* promoter. sgRNA-ribozyme expressed under *cdna1* promoter, terminated by *cbh2t* terminator.

A construct to continuously express a *Yarrowia lipolytica* codon optimized *cas9* in *T. reesei* was provided by VTT (pCPL23). A plasmid containing a *T. reesei* codon optimized *cas9* gene was given to our laboratory by Prof. Zou's laboratory (Liu, Chen et al. 2015). From this, the gene was copied with PCR and appropriate flanks were added to the primers' 3' ends to allow homologous recombination and insertion into the base vector pCPL34. The recombination was done by transforming the *cas9* PCR product and an *Ascl* & *Sml* restricted pCPL34 into *S. cerevisiae* ATCC 90845. Screening for successful assembly was done with restriction analysis and sequencing. The promoter was switched from *tef1* to *gpd1* with the same method. A nuclear localisation signal PKKKRKV is present at the C-terminus of Cas9 in all *cas9* expression constructs. Figure 10 collects the different *cas9* expression constructs used in this study.

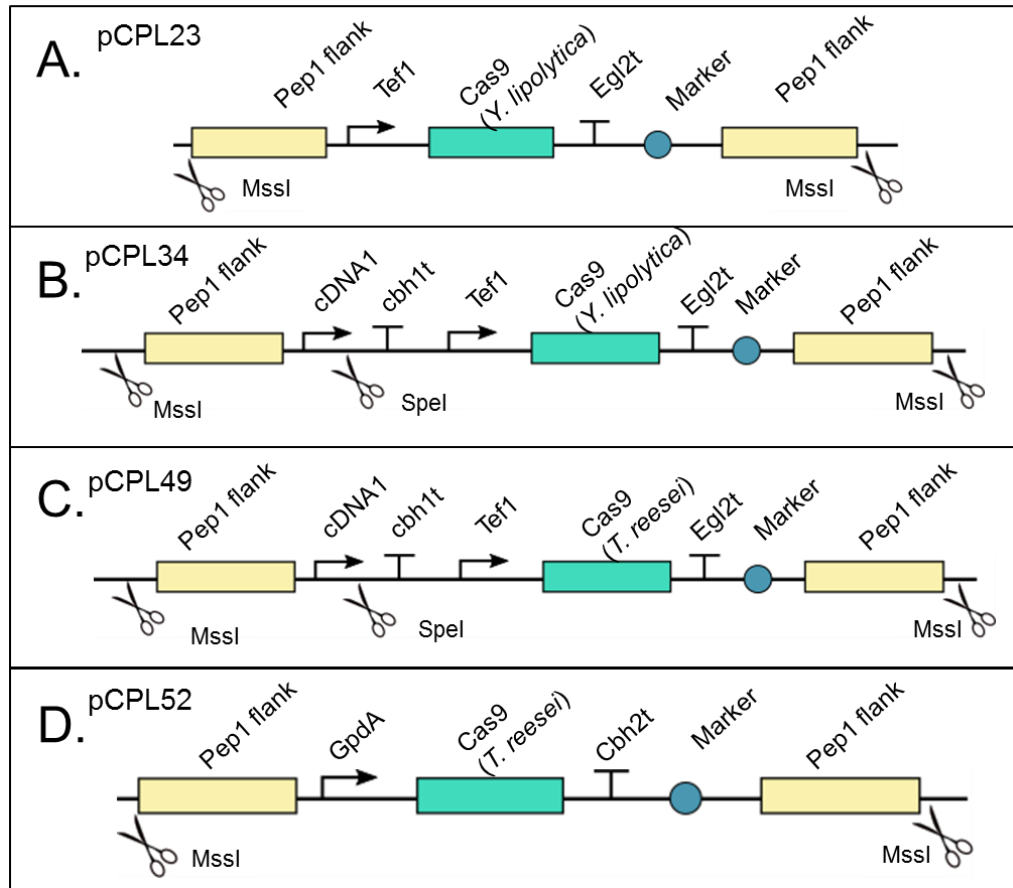


Figure 10 Base vector designs for *in vivo cas9* expression. Integration into *pep1* locus was directed by 1000 bp homologous flanks. **A.** pCPL23: *Y. lipolytica* optimized *cas9* expressed under *tef1* promoter, terminated by *egl2t* terminator. Both hygromycin and *pyr4* markers used. **B.** pCPL34: *Y. lipolytica* optimized *cas9* expressed under *tef1* promoter, terminated by *egl2t* terminator. *cDNA1* promoter and *cbh1t* terminator located upstream of *tef1* to allow for the insertion of an sgRNA-ribozyme complex. Both hygromycin and *pyr4* markers used. **C.** pCPL49: *T. reesei* optimized *cas9* expressed under *tef1* promoter, terminated by *egl2t* terminator. *cDNA1* promoter and *cbh1t* terminator located upstream of *tef1* to allow for the insertion of an sgRNA-ribozyme complex. Both hygromycin and *pyr4* markers used; **D.** pCPL52: *T. reesei* optimized *cas9* expressed under *gpdA* promoter, terminated by *cbh2t* terminator. *Pyr4* marker.

pCPL34 and pCPL49 functioned as base vectors and allowed for sgRNA-ribozyme cassette integration. This feature is missing from pCPL23 and pCPL52. pCPL53 is a replicate of pCPL52, but a green fluorescent protein (GFP) is fused to the N-terminus of Cas9.

4.8.2 gRNA Design and Production

Altogether four different target sites were used in this study, two for targeting *cp129* and two for targeting *crt1*. The target site selection and the methods for producing the gRNAs are described below. In this study, both sgRNAs and tracr:crRNA complexes were used. The same target sites were always used, regardless of the gRNA production method.

4.8.2.1 Target Site Design

The gRNA target sites were identified with Geneious CRISPR site finder. The program finds CRISPR target sites from a given sequence and compares these sites with the organism's whole genome sequence and gives a calculation of the sites' "CRISPR score". This score gives a good idea of the site's specificity in the organism. A score of 100% means that the program has not recognized any off-target sites with the given settings. The following settings to find target sites were used: 20 bp target site, NGG PAM-site, score all sites with *T. reesei* whole genome as the off-target database (Joint Genome Institute (JGI), *Trichoderma reesei* v2.0), maximum 3 bp mismatch against off-targets, maximum 1 bp mismatch allowed to be indels. The CRISPR score was the most important factor in choosing the target site, but also sites close to the GOI ORF's 5' and 3' ends were preferred.

In addition to using this program, the target sites were individually blasted against *T. reesei*'s complete genome and the specificity was checked by hand. Once the gRNA's target sequence had been chosen, an 80 bp tracrRNA tail (GTTTTAGAGCTAGAAATAGCAAGTTAAAT AAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTT) (Gao, Zhao 2014) was added to create an sgRNA. This tail associates with Cas9 protein.

The secondary structures of the gRNA's were predicted with a web-based RNA structure predictor (<http://rna.urmc.rochester.edu/RNAstructureWeb/>). Both the secondary structure of the whole gRNA and the target region alone were checked to make sure a secondary structure wouldn't interfere with the target site recognition.

4.8.2.2 gRNA Production

For the purpose of producing the sgRNAs *in vitro*, two approaches were used. A T7 promoter sequence (GTTTAACTAATACGACTCACTATA) was placed upstream of the sgRNA and a T7 terminator (GCGGCCGCGTTTAAAC) downstream. 30 bp flanking homology regions to pRS426 backbone vector were added towards the 5' and 3' regions of these expression sequences and the whole sequence was ordered as a gBlock (Integrated DNA Technologies, USA). The gBlock was then cloned into the pRS426 backbone through yeast-mediated homologous recombination and further into *E. coli* TOP10 cells. The plasmids were named pCPL39-40. 1 µg of the purified plasmid was linearized with NotI and the restriction reaction buffer was removed by QiaQuick PCR purification kit (Qiagen, Netherlands). 100 ng of the linearized, purified vector was used as a template for MegaScript RNA synthesis kit (ThermoFisher, USA). Transcription incubation times of 6 hours and 16 hours were used. The template DNA was removed with TURBO DNase I treatment and the remaining RNA was purified with the RNA purification kit (ThermoFisher, USA). The RNA concentration was determined with NanoDrop 2000 (ThermoFisher, USA). RNA was stored in -70 °C. This *in vitro* transcription method was only used for *crt1* sgRNAs.

The second approach used the GeneArt Precision gRNA Synthesis Kit (ThermoFisher, USA) and is based on a PCR-mediated template. This method is based on two universal primers and two primers that are specifically designed for each sgRNA. The universal primers include the sequence for a T7 promoter (TAATACGACTCACTATAG) and the tracrRNA sequence (Cas9 associating sequence) (GTTTTAGAGCTAGAAATAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTT). The forward sgRNA primer includes the T7 sequence and the first 16 bp of the target. The reverse primer includes the first 15 bp of the tracrRNA and the last 19 bp of the target. The sgRNA primers are first annealed into a primer-dimer template, which is then amplified by the universal primers (Figure 11).

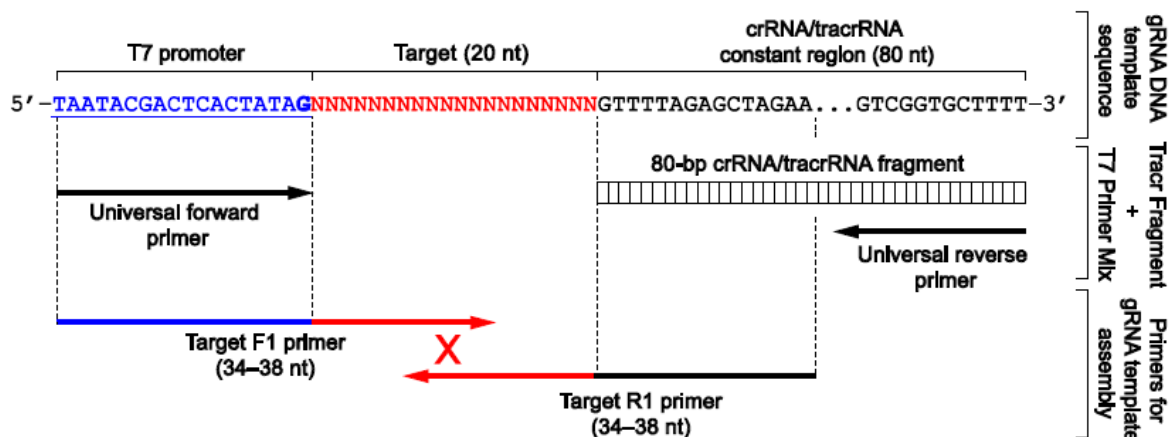


Figure 11 GeneArt Precision gRNA Synthesis Kit principle (www.thermofisher.com product #A29377).

The synthesized sgRNA was DNaseI treated and purified with the RNA purification kit (ThermoFisher, USA). Concentration was determined with NanoDrop 2000 (ThermoFisher, USA) and RNA was stored as 10 µg aliquots diluted in RNase free water in -70 °C.

The commercial tracrRNA and crRNAs were ordered from Integrated DNA Technologies (USA). The tracrRNA is a universal RNA molecule and crRNAs are designed specifically for each target. Both RNA's were diluted to 100mM final concentration in nuclease-free duplex buffer (Integrated DNA Technologies, USA) and stored in -20 °C.

4.8.2.3 sgRNA-ribozyme-complex

For the sgRNA-ribozyme-complexes, a 43 bp hammerhead ribozyme (NNNNNNCTGATGAGTC CGTGAGGACGAAACGAGTAAGCTCGTC) was added to the 5' end of the readily designed sgRNA (Gao, Zhao 2014). The first 5 nucleotides of this ribozyme sequence were altered specifically for each sgRNA to allow for base pairing between the 5' ends of the ribozyme and the 5' end of the sgRNA. In the 3' end of the gRNA, a constant 63 bp ribozyme sequence (GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGG AC) was used (Gao, Zhao 2014). The secondary structures were also predicted for these sgRNA-ribozyme complexes as described above in Section 4.8.2.1 Target Site Design. The sgRNA-ribozyme complex designs were done to target *crt1* only.

4.8.3 dDNA Design

Multiple types of donor DNAs (dDNAs) were used in this study. A single stranded oligonucleotide dDNA (ssdDNA) consisted of 15-30 bp long homologous flanks to the gene of interest. Between the flanks, a 20 bp primer sequence (CCTGTCAACTATCCCTACTC) was added. This primer sequence was later on used as a method of screening, together with an appropriate reverse primer attaching to the area of interest. This dDNA is later referred to as ssdDNA (Figure 12A). 2,5 µg of the ssdDNA diluted in TE buffer was used in each transformation. The ssdDNAs were ordered from Sigma-Aldrich (USA).

Another donor was a double stranded dDNA (dsdDNA) with 250-400 bp long homologous flanks to the gene of interest. The double stranded donor DNAs also included a primer sequence (CCTGTCAACTATCCCTACTC) between the flanks to be used for screening. This dDNA is later on referred to as dsdDNA (Figure 12B). 12 µg of dsdDNA digested with EcoRI from a TOPO plasmid was used in transformations with no further purification steps. The dsdDNAs were ordered as gBlocks from Integrated DNA Technologies (USA).

Finally a double stranded dDNA with a selection marker instead of a primer sequence was used. This dDNA included 1000-1500 bp long homologous flanks around the gene of interest. Instead of having a primer sequence between these flanks, a selection marker under an appropriate promoter was included. Both *amdS* and *pyr4* markers were used. This allowed for simple selection of positive clones. This dDNA is later on referred to as marker dDNA (Figure 12C). 10 µg of the marker dDNA was treated as traditionally transformed fragment prior to transformation (see Section 4.4). The marker dDNAs were assembled as traditional transformation cassettes as described in Section 4.4 *T. reesei* Transformation.

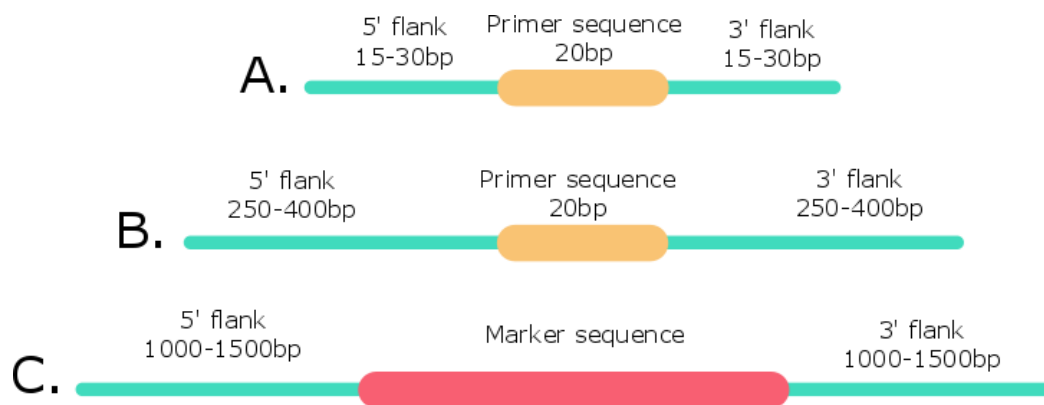


Figure 12 Different dDNAs used in this study. **A.** ssdDNA oligonucleotide containing 15-30 bp long homologous arms and a primer sequence for screening purposes. **B.** dsdDNA with 250-400 bp homologous arms and a primer sequence for screening purposes. **C.** Marker dDNA with 1000-1500 bp homologous arms and a selection marker (*amdS*, or *pyr4*) for selection and screening purposes.

4.8.4 Protein Transformation

Two different set-ups were used for protein transformations:

In the first set up, Cas9 protein was mixed with 30 μ l 10 \times Cas9 activity buffer, 35 μ l 2 \times STC buffer, and 1 μ g of IVT sgRNA as done by Pohl et al. (Pohl, Kiel et al. 2016). This mixture was incubated at +37 $^{\circ}$ C for 15 minutes before adding it to protoplasts. A dDNA was added into 250 μ l of protoplasts ($1-5 \times 10^6$ protoplasts/ml) directly after the Cas9 mixture followed by PEG-mediated protoplast transformation method (described in Section 4.4 *T. reesei* Transformation). This protocol was tested with 110 nM & 220 nM Cas9 protein concentrations. Two gRNAs (1 μ g of each) targeting the same gene ORF were used. This protein transformation method is later referred to as the Nygård method.

The second protein transformation method was based on Integrated DNA technologies' (USA) commercial Alt-RTM CRISPR-Cas9 system (CRISPR-Cas9 Genome Editing. 2016). In this method all components except the dDNA & 10 \times Cas9 activity buffer were ordered from Integrated DNA technologies (USA). The method uses tracr:crRNA complex instead of an sgRNA. 2 μ M of tracrRNA and 1 μ M of each crRNA in nuclease-free duplex buffer (Integrated DNA Technologies, USA) were incubated at 95 $^{\circ}$ C for 5 min and cooled on bench top. This tracr:crRNA mixture was diluted 1:2 in nuclease-free duplex buffer to acquire a 1 μ M

tracr:crRNA stock solution. Cas9 protein was diluted to 1,5 μ M stock concentration in 1 \times Cas9 activity buffer (diluted from 10 \times with RNase free H₂O). The components were mixed 1:1:1 (RNA: Cas9: 1 \times Cas9 activity buffer) to a total volume of 99 μ l and the mixture was incubated on bench top for 5 min. 250 μ l of protoplasts ($1-5 \times 10^6$ protoplasts/ml) were added and directly after, the dDNA was added, followed by protoplast transformation protocol. This method is later referred to as the IDT (Integrated DNA Technologies) method.

4.8.5 AMA1 Plasmid Transformation

An AMA1 plasmid expressing *cas9* under a *tef1* promoter was provided by Prof. Mortensen's laboratory (Nodvig, Nielsen et al. 2015). The plasmid had previously been used in *Aspergillus niger* and the plasmid includes an expression construct for a sgRNA and *cas9* as well as AMA1 replication site and a fungal marker. We used an AMA1 plasmid with a hygromycin marker under a *trpC* promoter, which is known to function in *T. reesei*. We used the AMA1 plasmid to express the *cas9* only and no sgRNA sequence was cloned to the sgRNA expression site. 1 μ g of the AMA1 plasmid was transformed into 250 μ l of protoplasts ($1-5 \times 10^6$ protoplasts/ml). 0,5 μ g of each IVT sgRNAs targeting *crt1* were added to the mixture as well as 10 μ g of the desired donor. AMA1 transformations were done using an ssdDNA, dsdDNA, and a marker dDNA with either *pyr4* or *amdS* marker. Successful transformants were selected on hygromycin plates for AMA1 selection and a desired selection for dDNA. After transformation, the clones were streaked on TrMM-triton plates with no hygromycin selection, as the AMA1 plasmid is often lost from the cells in an efficient and quick manner. DNA was collected from the streaks and the transformants were genotyped with Phire Plant Direct PCR Kit (ThermoFisher, USA) for successful dDNA integration.

4.9 Transporter Deletions

Fragments from pCPL27, pCPL28, pCPL29, and pCPL50 deletion constructs were used to introduce transporter deletions into *T. reesei* M127. Introduction was done using the traditional *T. reesei* transformation method, because setting up the CRISPR/Cas9 method for all of them took too much time. However, $\Delta cpl29$ and $\Delta crt1$ were done as CRISPR/Cas9 deletions as well. All analysed strains were created by the traditional transformation method. *Pyr4* was used as a selection method for all deletions.

4.10 Enzyme Assays

Two enzyme assays were used to determine cellulase activities in *T. reesei* transporter deletion strains.

The Azo-barley glucan assay (Megazyme, Ireland) was used to determine the samples' endoglucanase activities. The assay is based on a polymerized substrate that is insoluble in the assay's precipitant solution. Endoglucanases depolymerise the substrate into a soluble form. The assay was done in +40 °C heat block. Pre-heated Azo-barley glucan substrate was diluted 1:3 in Assay Buffer (100 mM NaAc buffer, pH 4.5, 500 mg/l BSA, 200 mg/l NaN₃) and added to 1,5 ml eppendorf tubes as 100 µl aliquots and placed into +40 °C heat block for ≥ 5 min. The samples were diluted accordingly and pre-heated to +40 °C. 10 µl of sample was added to the substrate, vortexed and incubated exactly 10 min in +40 °C. After the incubation, 300 µl of Precipitant Solution B (30 g/l NaAc, 4.0 g/l zinc acetate diluted in 300 ml H₂O, pH 5.0, mixed with 700 ml EtOH (total volume 1000 ml)) was added to the tube and vortexed. After 5min in room temperature, the samples were vortexed again and centrifuged at 1000 g for 10 min. 200 µl of the supernatant was placed to microtiter plate and the 590 nm absorbance was measured with Varioskan Flash (ThermoFisher, USA). The absorbance values were analysed according to standard curve 4 provided by Megazyme (<https://secure.megazyme.com/Azo-Barley-Glucan>).

The MULac assay measures β -glucosidase and other cellulase activities. The assay is based on a fluorescent substrate, normally fluorescing at 356 nm, but at 445 nm after enzymatic

degradation. Standard samples containing 1,25 μ M – 40 μ M 4-methylumbelliferone (M1381-25G, Sigma) in 50 mM NaAc (pH 5.0) were prepared in duplicates. Each sample was diluted in 50 mM NaAc (pH 5.0) appropriately and prepared in triplicates. The substrate 4-methylumbelliferyl β -D-lactopyranoside (M2405-100MG) was diluted into 2 mM working solution in 50 mM NaAc (pH 5.0). The substrate and sample (/standard) were mixed 1:1 in a black 96-well plate and incubated at room temperature for 15 min. The reaction is stopped by adding 100 μ l of 1 M Na₂CO₃ and the fluorescent signal is measured with Varioskan Flash (Thermo Scientific, USA). A standard curve is made based on the fluorescent signal from standard samples and each sample's enzyme concentration is determined based on this.

4.11 Sequencing

All sequencing was done with at GATC-Biotech (Germany) using Lightrun. A new primer was used for every ~500 bp of sequence. Primers were provided to the service provider along with the template DNA for sequencing.

5 Results

5.1 Plasmids

Table 6 collects the plasmids that were constructed during this study.

Table 6 Plasmids constructed during this study.

Plasmid name	Contents	Bacterial selection	<i>T. reesei</i> selection
pCPL34	<i>Y. lipolytica</i> optimized <i>ptef1-cas9</i> -NLS base vector	Ampicillin	Hygromycin
pCPL39	<i>pt7-crt1</i> sgRNAv1 in pRS426	Ampicillin	-
pCPL40	<i>pt7-crt1</i> sgRNAv2 in pRS426	Ampicillin	-
pCPL41	<i>Crt1</i> dsdDNA in TOPO	Kanamycin	-
pCPL42	<i>Crt1</i> sgRNAv1 in pCPL34	Ampicillin	Hygromycin
pCPL43	<i>Crt1</i> sgRNAv2 in pCPL34	Ampicillin	Hygromycin
pCPL47	<i>ptef1-trecas9</i> -NLS in pCPL34	Ampicillin	Hygromycin
pCPL48	<i>Crt1</i> deletion cassette in pRS426	Ampicillin	AmdS
pCPL49	<i>ptef1-trecas9</i> -NLS in pCPL34	Ampicillin	Pyr4
pCPL50	<i>Crt1</i> deletion cassette in pRS426	Ampicillin	Pyr4
pCPL52	<i>pgpdA-trecas9</i> -NLS in pRS426	Ampicillin	Pyr4
pCPL53	<i>pgpdA-trecas9</i> -GFP-NLS in pRS426	Ampicillin	Pyr4

5.2 gRNA Design

Two gRNAs were chosen for each target gene. The chosen *crt1* gRNAs have no more than a 3 bp unspecific match in the 3' end of the protospacer sequence. For *cp/29* gRNA1 and 2 a 3 bp and 10 bp match in the 3' end to an off-target were found. According to review literature, at least 8-12 corresponding base pairs in the 3' end are needed to accurately specify the target site (Sander, Joung 2014). The chosen target sites are located no more than 500 bp from the 5' and 3' ends of the GOI's ORF.

The structure predictions revealed no significant target sequence secondary structures. 20 possible secondary structures for both *crt1* sgRNAs and 6 and 2 for *cp/29* sgRNAs 1 and 2, respectively, were found. In these whole sgRNA secondary structures, the likelihood of a secondary structure affecting the target sequence was moderate or low. The likelihood of a secondary structure affecting the target sequence of the sgRNA-ribozyme-complexes was also moderate to low.

Table 7 summarizes the target sequences used for each targeted gene and their CRISPR score as calculated by Geneious CRISPR site finder.

Table 7 gRNA target sequences. Two target sites were chosen for each gene.

Target	Target sequence	CRISPR Score
<i>Crt1</i> -1	GCGCCATGTACAACACTCT	100,00%
<i>Crt1</i> -2	CGAGACCAAGGGACGCACTT	100,00%
<i>Cp/29</i> -1	Proprietary - not shown	99,63%
<i>Cp/29</i> -2	Proprietary - not shown	99,74%

5.3 gRNA Synthesis

MegaScript gRNA IVT synthesis kit (ThermoFisher, USA) resulted in yields of ~4 000 ng per reaction. GeneArt precision gRNA synthesis kit (ThermoFisher, USA) proved to be much more efficient for the IVT gRNA production providing yields of 50 000 ng per reaction. This was likely due to a smaller amount of template DNA in the Megascript synthesis. An overnight incubation of the MegaScript transcription mixture resulted in even smaller yields of RNA. This may be due to RNase contamination in the reaction mixture and subsequent RNA degradation over a long incubation period.

5.4 *T. reesei* Continuous *Cas9* Expression & Ribozyme Constructs

After the assembly of pCPL49, *T. reesei* codon optimized *cas9* expression cassette, sequencing analysis revealed that some base pairs in the 5' end of the coding region of *cas9* remained as *Y. lipolytica* codon optimized sequence. This was likely due to contamination. A clone with 69 bp of *Y. lipolytica* codon optimized bases in the beginning of the ORF was chosen as our *T. reesei* optimized *cas9*.

pCPL23 (*Y. Lipolytica* codon optimized *cas9*, hygromycin marker), pCPL26 (*Y. lipolytica* codon optimized *cas9*, *pyr4* marker), and pCPL49 (*T. reesei* codon optimized *cas9*, *pyr4* marker) were transformed into M124, M127 & M1692 according to the selection marker in use. Successful transformation was determined by PCR for *cas9* ORF, *pep1* ORF, and 5' and 3' integrations. From initial screening, four positive M127-cpl49 clones were chosen to be purified. Only one of these grew when single spores were plated and all single spores gave a negative *cas9* ORF signal. No positive clones of M127 cpl49 were found. After cleaning, two PCR positive M124-cpl23, M1692-cpl26, and M1692-cpl49 clones were analysed with a western blot and qPCR. See Table 8 for qPCR cycle counts.

A transcript was detected in both M1692 cpl26 clones, but M124 cpl23 and M1692 cpl49 clones gave negative signals. M124 cpl23 is likely a false positive clone, and further PCR screens could maybe prove that the expression cassette has been lost from the strain. The negative signal of M1692 cpl49 could be due to a lost cassette or an issue with qPCR primers.

The C144-C145 primers create a ~900 bp product, which may be too long for detection under the set qPCR conditions. Despite the negative qPCR signals, the clones were analysed with a western blot as well. Figures 13 and 14 present the images of Cas9 detection blots.

Table 8 qPCR analysis for *cas9* transcription in M124 and M1692. M717 is a *pyr4*⁺ control strain of M1692.

Sample	PP181-PP182 (<i>Y. lipolytica cas9</i>) (C _q)	C144-C145 (<i>T. reesei cas9</i>) (C _q)	<i>Gpd1</i> (housekeeping) (C _q)
M124	33,93	n/a	34,42
M717	34,67	n/a	21,31
M124-cpl23-36-1	34,41	n/a	21,01
M124-cpl23-4-9	33,75	n/a	19,30
M1692-cpl26-2-1	16,69	n/a	20,78
M1692-cpl26-2-2	18,90	n/a	19,75
M1692-cpl49-16-5	n/a	33,61	19,80
M1692-cpl49-16-6	n/a	36,61	22,07

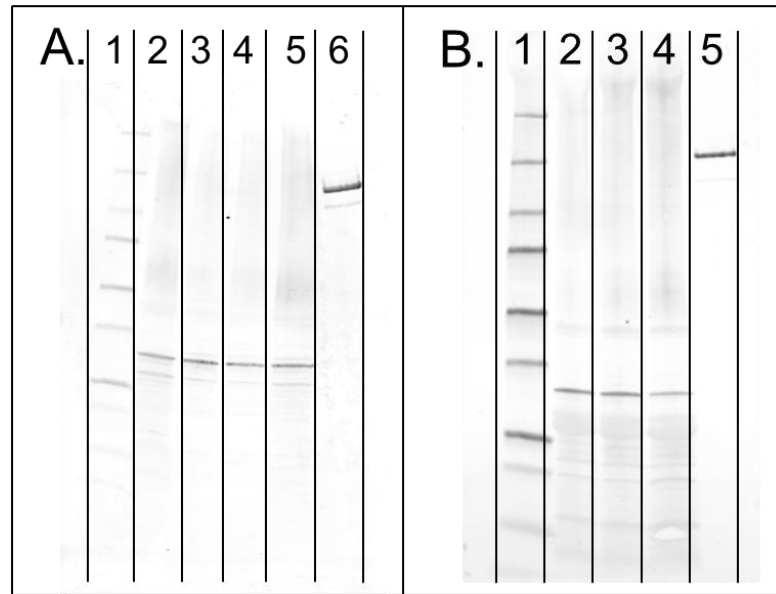


Figure 13 Western blots for Cas9 detection from protein extracts. **A.** 1. Marker; 2. M124; 3. M124-cpl23-4-9; 4. M124-cpl23-36-1; 5. M124-cpl23-36-3; 6. 200ng Cas9. **B.** 1. Marker; 2. M1692-pCPL49-16-5; 3. M1692-pCPL49-16-6; 4. M1692-pCPL49-9; 5. 200ng Cas9 protein

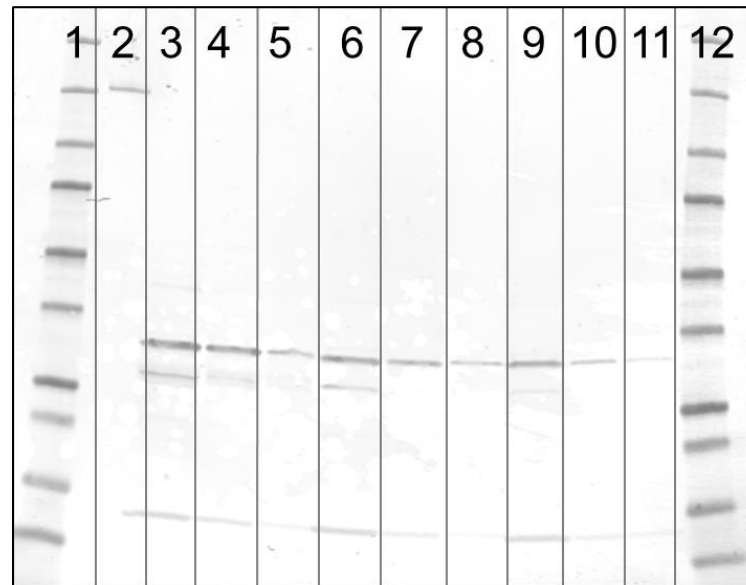


Figure 14 A western blot for Cas9 detection from protein extracts. 1. Marker; 2. 100ng Cas9 protein; 3. M717-26-2-1 1:3 diluted; 4. M717-26-2-1 1:6 diluted; 5. M717-26-2-1 1:12 diluted; 6. M717 26-2-2 1:3 diluted; 7. M717-26-2-2 1:6 diluted; 8. M717-26-2-2 1:12 diluted; 9. M124 1:3 diluted; 10. M124 1:6 diluted; 11. M124 1:12 diluted; 12. Marker.

Cas9 protein was not detected from any of the clones. The western blot does detect something of a much smaller size than the whole Cas9 protein. This however doesn't seem to be a Cas9 degradation product, but rather some background from the whole protein extract itself as the same pattern is detected from M124 parent strain without a *cas9* expression construct.

pCPL52 and pCPL53, *T. reesei* optimized *cas9* under *gpdA* promoter with and without GFP fusion, respectively, were transformed into M127. Only 10 and 1 transformants with pCPL52 and pCPL53, respectively, grew through the transformation plate agar. None of these transformants gave a positive PCR signal for *T. reesei* optimized *cas9* ORF. No GFP signal was detected from pCPL53 transformation plates. This suggests that the *gpdA* promoter is too strong for continuous *cas9* expression in *T. reesei* and that at certain concentrations Cas9 is toxic for *T. reesei*. The small colonies that do not grow through the transformation plate agar (Figure 15) may be background or the clones in which the *cas9* expression cassette has been integrated correctly and are thus lagged in growth.

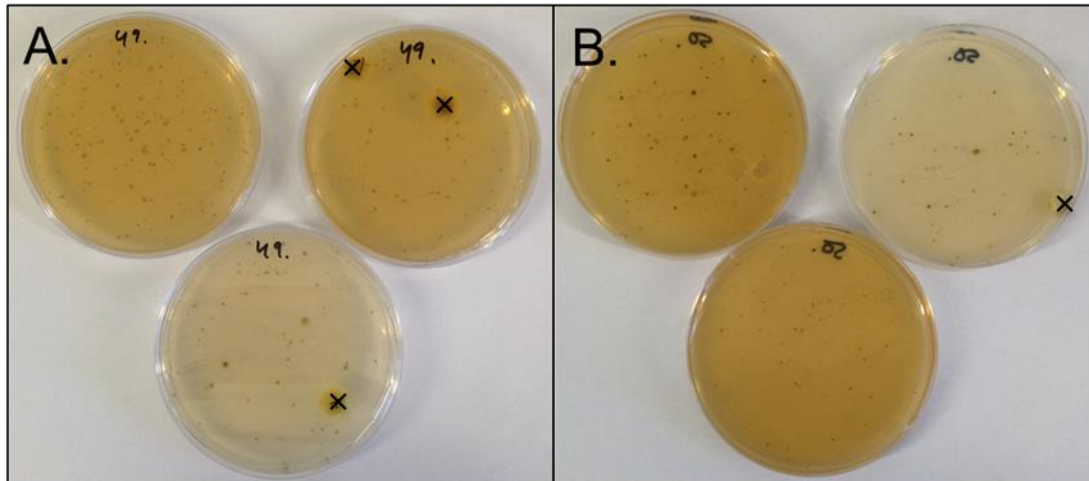


Figure 15 Transformation plates of pCPL52 and pCPL53 in M127 12 days post-transformation **A.** pCPL52 in M127. Marked colonies were checked for *T. reesei* optimized *cas9* ORF by PCR; **B.** pCPL53 in M127. Marked colonies were checked by PCR for *cas9* ORF.

These results suggest that *cas9* has not successfully been expressed in any of our transformants. This could be due to the promoters not being suitable. *Cas9* under *gpdA* promoter has been reported to be toxic in *Penicillium* (Y. Nygård, personal communication)

and the situation may be similar in *T. reesei*. It's however difficult to know if the expression cassette is working with the *gpdA* promoter, or if something other than *cas9* is causing the growth issue.

With *tef1* it is not certain whether the promoter is too strong, too weak or just right. A clear *cas9* transcript was detected by qPCR in M717 cpl26 strains, but the protein was not detected in western blots. Regarding the qPCR, very small amounts of mRNA should be sufficient for translation and very small amounts of Cas9 protein should be detectable by a western blot. As a mixture of dT₁₈ and random primers was used for cDNA synthesis, it may be possible that the transcript is truncated and not processed into a full mRNA. In this case the random primers may still be transcribing the transcript into cDNA, but it would either never be processed into a protein or would create a truncated protein perhaps unrecognizable for the primary antibody in use. Another possibility is that the transcript is produced as expected, but the translation initiation doesn't take place for some reason or that the translation is prematurely terminated. One option is also that the Cas9 protein is very toxic in *T. reesei* and is rapidly degraded after translation. This could keep the Cas9 concentrations too low for western blot detection.

For M124-cpl23 and M1692-cpl49 the negative western blot results are in line with the qPCR results. In this case the expression cassette seems to simply not be functioning. It maybe incorrectly integrated, lost from the strain or a similar problem of transcript truncation takes place as hypothesized for M717-cpl26 strains, but so early that the transcript is not reverse transcribed into cDNA. What's notable is that the only difference between pCPL26 and pCPL23 is the fungal marker in use, but they are giving different qPCR signals. One possibility is, of course, that the qPCR simply didn't work for some reason for M124-cpl23 samples. This is unlikely though, since the qPCR was done for two clones and triplicate samples for each were used. pCPL49 expression cassette on the other hand includes an sgRNA-ribozyme expression system upstream of the *cas9* expression cassette. There is a possibility that the transcript from the *cdna1* promoter is not successfully terminated by the *cbh2* terminator and creates a transcript that's not recognizable to the translation initiation machinery. This possibility doesn't however explain what may have happened in M717-cpl26 and M124-cpl23 expression strains.

Transformants containing an sgRNA-ribozyme and *cas9* expression systems (pCPL42 & pCPL43) were screened according to table 9.

Table 9 PCR screening results of pCPL42 and pCPL43 transformed into M124. Hygromycin selection.

	dDNA ⁺	<i>Crt1</i> ORF ⁻	<i>Cas9</i> ORF ⁺	<i>Pep1</i> ORF ⁻	<i>Cas9</i> 5' ⁺	<i>Cas9</i> 3' ⁺
M124-cpl42, ssdDNA C112	0/47	0/10	10/10	2/10	2/10	2/10
M124-cpl42, dsdDNA pCPL41	0/48	0/10	10/10	2/10	2/10	2/10
M124-cpl43, dsdDNA pCPL41	0/45	0/10	10/10	0/10	0/10	1/10
M124-cpl42, pCPL50 marker dDNA	n/a	2/70 (3%)	51/65	n/a	n/a	n/a
M124-cpl42, pCPL50 marker dDNA	n/a	4/69 (6%)	49/65	n/a	n/a	n/a

The efficiency to disrupt *crt1* was very low. When an oligo ssdDNA (C112) and a dsdDNA (pCPL41) without marker sequences were used no negative *crt1* ORFs were detected. All analysed clones did however give a positive *cas9* ORF signal and in four cases the sgRNA-ribozyme-*cas9* expression cassette was integrated correctly. If both the *cas9* and the sgRNA were expressed correctly, *crt1* should be disrupted or the cells should die from being unable to repair the DSB created by CRISPR/Cas9. To assess whether the CRISPR/Cas9 components were expressed, a qPCR analysis was done to detect *cas9* and sgRNA transcripts. The *cas9* detection was done for PCR positive M127-cpl42-cpl41-1 and M127-cpl42-cpl41-6 with dT₁₈ primers. No transcript was detected in either strain (data not shown). The qPCR was repeated for M127 cpl42-cpl41-6 with dT₁₈ and random primer mix. See Table 10 for *cas9* qPCR cycle counts.

Table 10 qPCR cycle counts for M127-cpl42-cpl41-6 sgRNA-ribozyme-*cas9* expression strain.

Sample	PP181-PP182 (<i>Y. lipolytica</i> opt. <i>Cas9</i>) (C_q)	<i>Gpd1</i> (housekeeping) (C_q)
M127-cpl42-cpl41-6	21,73	20,92
M124	33,93	34,42

Cas9 transcript level was normalized to *gpd1* house-keeping gene with relative quantification and revealed an expression level half of *gpd1*'s. It's a low expression level, but the transcript is detected at a much higher level than in the M124 control strain. To analyse the sgRNA expression, RNA from M127-cpl42-cpl41-1 and M127-cpl42-cpl41-6 clones were analysed along with plain IVT *crt1* sgRNA. The cDNA synthesis was done with a designated primer C145 instead of dT₁₈ primers. No sgRNA transcript was detected in either the total RNA or a plain IVT *crt1* sgRNA sample (data not shown). This may be due to a problem with the primers, which would interfere with both the cDNA synthesis and qPCR amplification. It's also possible that the sgRNA has such strong secondary structures that it was not detected with the reaction cycle in use. A western blot was done to detect Cas9 from mycelia samples and protein extracts. No Cas9 protein was detected (Figures 16 & 17).

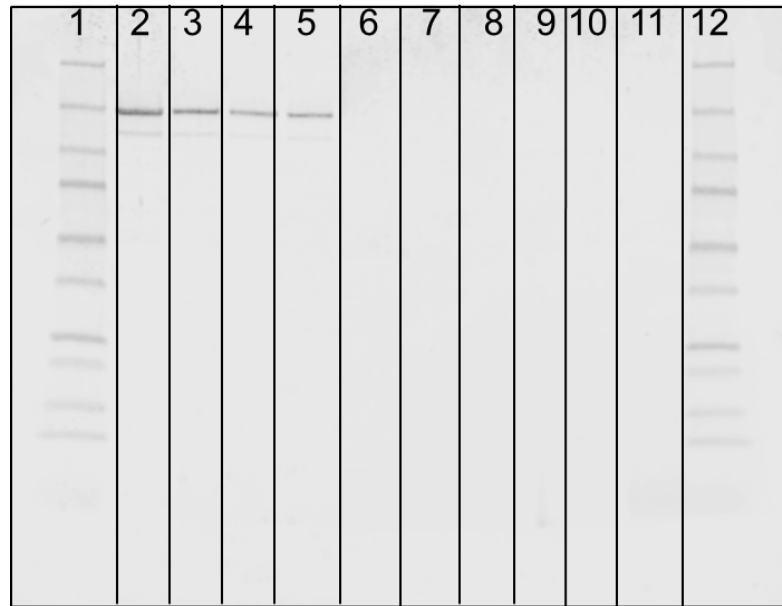


Figure 16 A western blot for Cas9 in M124-cpl42-cpl41 mycelia samples. **1.** Marker **2.** 400ng Cas9 **3.** 200ng Cas9 **4.** 100ng Cas9 **5.** 50ng Cas9 **6.** M124-cpl42-cpl41-1 1:3 diluted **7.** M124-cpl42-cpl41-1 1:6 diluted **8.** M124-cpl42-cpl41-1 1:12 diluted **9.** M124-cpl42-cpl41-6 1:3 diluted **10.** M124-cpl42-cpl41-6 1:6 diluted **11.** M124-cpl42-cpl41-6 1:12 diluted **12.** Marker.

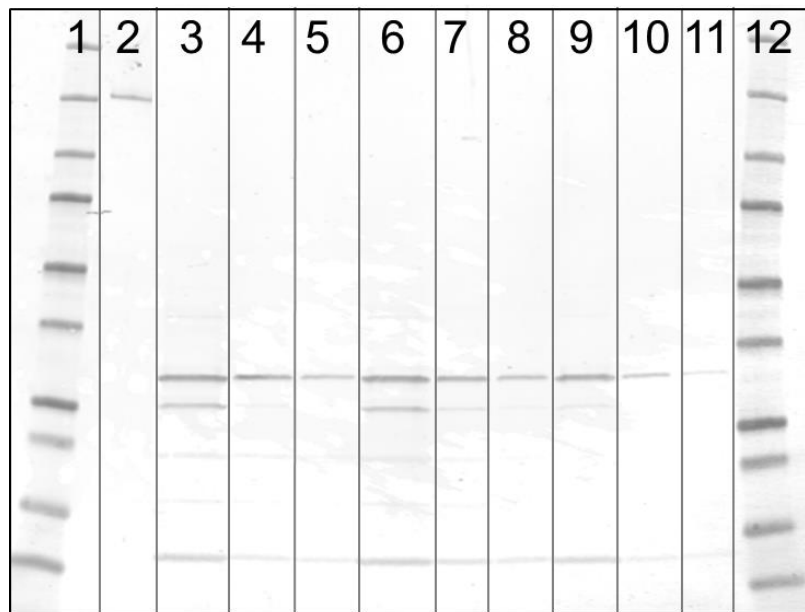


Figure 17 A western blot for Cas9 detection in M124-cpl42-41 protein extracts. **1.** Marker; **2.** 100ng Cas9 protein; **3.** M124-cpl42-cpl41-1 1:3 diluted; **4.** M124-cpl42-cpl41-1 1:6 diluted; **5.** M124-cpl42-cpl41-1 1:12 diluted; **6.** M124-cpl42-cpl41-6 1:3 diluted; **7.** M124-cpl42-cpl41-6 1:6 diluted; **8.** M124-cpl42-cpl41-6 1:12 diluted; **9.** M124 1:3; **10.** M124 1:6 diluted; **11.** M124 1:12 diluted; **12.** Marker.

Since the efficiency of *crt1* deletions was very low when using a marker dDNA pCPL50 and *cas9* ORF was detected in most of the clones, no further analysis was done for these transformants. The problem most likely lies in the inefficient translation of *cas9* and possible absence of the sgRNA altogether. Another project was going on simultaneously using the same approach of expressing the sgRNA between ribozyme sequences followed by *cas9* under *tef1* promoter and was also found unsuccessful. In their case some sgRNA's were detected with qPCR but *cas9* translation product was undetectable in a western blot (data not shown). This would suggest that the problem most likely lies in the translation of *cas9*. In this case, as with M1692-cpl49, it may be that the sgRNA transcript is not efficiently terminated by the *cbh2* terminator and compromises the translation initiation of *cas9*.

The different qPCR results when using dT₁₈ primers and a mixture of dT₁₈ primers and random primers support the hypothesis that a complete *cas9* transcript is not generated. However, the primer mixture is dependent on the cDNA synthesis kit, and the problem could also lie in the Transcriptor First Strand cDNA synthesis kit, although a *gpd1* transcript was detected when using this method.

All in all, further studies are called for to determine the cause of non-existing Cas9 protein in all of the expression strains. The results presented above are not thorough enough to determine whether the problem is on transcription or translation level. Repetitions of the qPCR analysis using different primers for reverse transcription and for *T. reesei* optimized Cas9 detection should be done.

5.5 AMA1 Transformation

The AMA1 plasmid with a *cas9* expression cassette was used to try to achieve transient *cas9* expression. The plasmid was transformed into M124 along with a dsdDNA. The transformation plates became so full that picking individual clones was difficult. Out of 50 screened, M124 dsdDNA and AMA1 transformed clones, none were positive for $\Delta crt1$ and no signal was obtained when the dsdDNA's primer sequence was used for amplification (data not shown). At this stage we moved on to only using marker dDNAs to add selection pressure. It seemed that the AMA1 transformation was so efficient that the hygromycin

selection alone was not sufficient to identify positive clones. Table 11 collects the results of the AMA1 transformations with a marker dDNA.

Table 11 AMA1 transformation results using a marker dDNA.

Strain	dDNA target	dDNA selection	AMA1 selection	ORF ⁺
M124	<i>Crt1</i>	AmdS	Hygromycin	0/58 (0%)
M124	<i>Crt1</i> , no gRNA ctrl	AmdS	Hygromycin	0/28 (0%)
M127	<i>Crt1</i>	Pyr4	Hygromycin	1/25 (4%)
M127	<i>Crt1</i> , no gRNA ctrl	Pyr4	Hygromycin	15/78 (19%)
M1692	<i>Crt1</i>	Pyr4	Hygromycin	31/104 (30%)
M1692	<i>Crt1</i> , no gRNA ctrl	Pyr4	Hygromycin	40/105 (38%)
M127	<i>Crt1</i>	Pyr4	Hygromycin	6/67 (9%)
M127	<i>Crt1</i> , no gRNA ctrl	Pyr4	Hygromycin	4/103 (4%)
M1692	<i>Cpl29</i>	Pyr4	Hygromycin	2/42 (5%)
M1692	<i>Cpl29</i> , no gRNA ctrl	Pyr4	Hygromycin	32/105 (30%)

From Table 11 we can conclude that the AMA1 transformation method did not provide a significant benefit to introducing *crt1* or *cpl29* deletions. When deleting *cpl29* the AMA1 actually significantly reduced the amount of $\Delta cpl29$ transformants. In this experiment the number of screened transformants is the number of colonies that grew through the agar. The transformation efficiency is thus low, and the results are statistically unreliable as so few clones were screened from some of the transformations. This would suggest that there's something in the AMA1 plasmid that is not very well tolerated by *T. reesei*. This could be a toxic effect of *cas9*, or something else. These results suggest that the *tef1* promoter introduces too strong of a *cas9* expression. Unfortunately there wasn't a chance to test an AMA1 plasmid with something other than *cas9* expression to determine the effect of other AMA1 components.

5.6 Cas9 Protein Transformation

An alternative approach was taken to avoid having to express *cas9* in the strain as this seemed difficult to achieve. Testing the feasibility of a transient protein transformation seemed like an attractive alternative. *Crt1* and *cp129* were targeted with two IVT RNAs along with marker dDNAs (pCPL50 & pCPL29) with the Nygård method. Final concentrations of 110 nM and 220 nM of Cas9 protein were tested (ThermoFisher, USA).

As Table 12 shows, there was no significant increase in removing the *crt1* ORF with CRISPR/Cas9 in M127. There was only a slight difference between the CRISPR/Cas9-method and the control reaction in which no gRNA was added. For the $\Delta crt1$ transformants, no further analysis was done based on the ORF screening.

Table 12 Screening results of Nygård method protein transformation in M127 (pyr4 selection in dDNA).

Target	[Cas9]	ORF ⁻	5' ⁺	3' ⁺	Total (all PCRs OK)
<i>Crt1</i>	110nM	28/95 (29%)	n/a	n/a	n/a
<i>Crt1</i>	220nM	22/75 (29%)	n/a	n/a	n/a
No gRNA ctrl (pCPL50 dDNA)	220nM	23/100 (23%)	n/a	n/a	n/a
<i>Cp129</i>	110nM	23/96 (24%)	18/19	15/19	12/96 (13%)
<i>Cp129</i>	220nM	26/94 (28%)	10/23	19/23	10/94 (11%)
No gRNA ctrl (pCPL29 dDNA)	110nM	26/46 (57%)	10/33	16/32	5/46 (11%)
No gRNA ctrl (pCPL29 dDNA)	220nM	22/34 (65%)	12/27	15/27	5/34 (15%)

For *cp129* it seemed that the deletion was more effective in the control reactions. To see if CRISPR/Cas9 had increased the efficiency of correct integration, both 5' and 3' integrations were checked. When clones with negative ORF and positive integrations were summed,

there was no significant difference between the CRISPR/Cas9 and control transformations. What's interesting is that the control reaction had a clear effect on the cells. Only 46 and 34 clones grew through the transformation agar when 110nM and 220nM Cas9 protein was transformed, respectively. Based on this, and the unexpected ORF screening results, the *cpI29* deletion was repeated with the Nygård method and the IDT protein transformation method. This time the Cas9 protein used in both methods came from a batch ordered from Integrated DNA Technologies (USA). Table 13 presents the screening results.

Table 13 Screening results of repeated Nygård method and IDT protein transformation method for *ΔcpI29* in M127.

Target	[Cas9]	ORF ⁻	5' ⁺	3' ⁺	Total (all PCRs OK)
<i>CpI29</i> (pCPL29 dDNA, Nygård method)	220nM	52/108 (48%)	66/108 (61%)	82/108 (76%)	34/108 (31%)
No gRNA ctrl (pCPL29 dDNA, Nygård method)	220nM	6/33 (18%)	6/33 (18%)	5/33 (15%)	0/33 (0%)
<i>CpI29</i> (pCPL29 dDNA, IDT method)	100nM	66/96 (69%)	63/96 (66%)	93/96 (97%)	43/96 (45%)

As can be seen from Table 13, this time there was a clear difference between the control reaction and the CRISPR/Cas9 transformation, in which at least 31% of clones were correct for all PCRs. This is roughly three times better than when the method was done for the first time. This could be due to the different protein batch. The previously used Cas9 from ThermoFisher had been thawed and used multiple times as a control for western blots, and may have lost some of its catalytic activity due to this. Again, only 33 clones grew for the control reaction. Notably, the IDT method seems to be working even better than the Nygård method. 97% of the clones gave a positive 3' signal. ORF is also absent in 66% of the clones, which is more than has been previously recorded. Also, the IDT method seemed to result in high transformation efficiencies. From plate images, we can see a clear difference between the transformation plates when using the Nygård method vs. the IDT method (Figure 18).

The IDT method was so efficient that it wasn't possible to count the colonies and determine the transformation efficiency at this point in time. The transformation mix would be more diluted in further experiments.

Based on these results, we discarded the Nygård method and moved on to using the IDT method alone, although it seems that the Nygård method produced fairly high amounts of correctly integrated donor cassettes (31%).

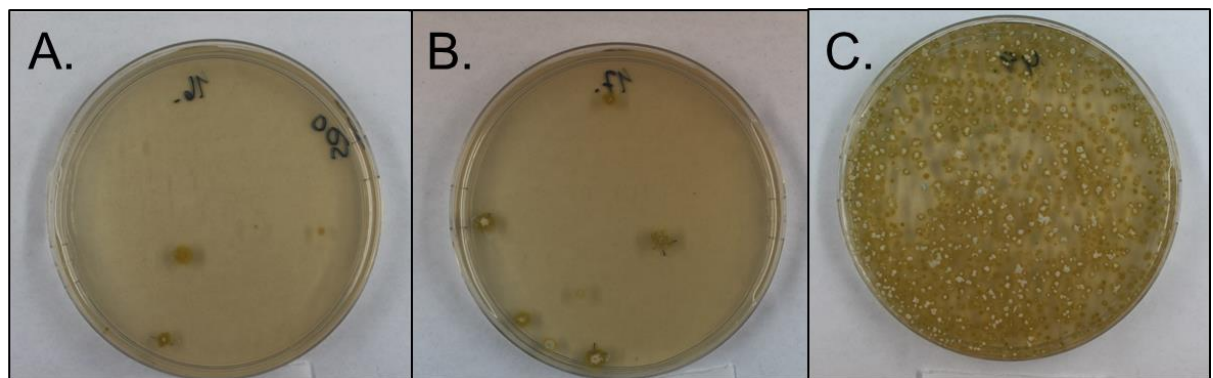


Figure 18 Transformation plates from Nygård & IDT methods 13 days post-transformation. 500 μ l protoplasts per plate. **A.** Nygård method no gRNA control in M127. **B.** Nygård method $\Delta cpl29$ M127. **C.** IDT method $\Delta cpl29$ M127.

In addition to repeating the IDT method for deleting *cpl29*, we also tested a deletion of *crt1* and a double deletion of both genes at once. As an effort to understand why so many cells die from Cas9 transformation without any RNA, we also tested all of these deletions in M123, which is a parent of M124 and contains *mus53⁺* (*pyr4⁻*). We hypothesized that Cas9 may be introducing unspecific cuts, and the lack of the nonhomologous repair gene *mus53* prevents the cells from repairing the DSBs, causing cell death. Table 14 collects PCR results from CRISPR/Cas9 transformation and controls in strains M127 and M123. The overall number of positive clones was determined by picking clones, which gave a negative ORF signal and positive signals for both integration amplifications. All three signals were analysed for all transformants. The number of analysed clones for each transformant was 96. If less than 96 clones were analysed, it's due to them not growing through the transformation plates.

Table 15 presents the number of colonies in each transformation plate. All transformations were done to the same batch of protoplasts, for M127 and M123 $7,5 \times 10^6$ and $8,25 \times 10^6$ protoplasts were transformed in each reaction, respectively. The number of positive clones is determined by *the number of colonies / plate \times total percentage of positive clones*, as shown in Table 14. The transformation efficiency calculation follows the formula *number of positive clones / μ g DNA / dilution*.

For each transformation, the number of colonies was counted from three plates. The colony number shown in Table 15 is an average of the three plates. Colony numbers were determined on days 9, 10, and 11 after transformation. Due to differences in growth speed and human error, the colony numbers may not be as accurate as possible. The number of screened colonies (Table 14) may give a better idea of how many colonies grew through the transformation plate. In the colony count, background clones that do not reach the surface of the transformation plate are not disregarded.

Table 14 PCR results of IDT CRISPR method in M127 and M123. A *pyr4* marker dDNA was used for each transformation.

Strain	Target	Method	ORF ⁻	5' ⁻	3' ⁻	Total (All PCRs OK)
M127	<i>Cpl29</i>	IDT	77/96 (80%)	72/96 (75%)	93/96 (97%)	56/96 (58%)
M127	<i>Cpl29</i>	Traditional	1/50 (2%)	17/50 (34%)	25/50 (50%)	1/50 (2%)
M127	<i>Cpl29</i>	IDT, no gRNA ctrl	50/68 (74%)	9/68 (13%)	10/68 (15%)	1/68 (1%)
M127	<i>Crt1</i>	IDT	95/96 (99%)	2/96 (2%)	3/96 (3%)	2/96 (2%)
M127	<i>Crt1</i>	Traditional	38/96 (40%)	1/96 (1%)	6/96 (6%)	0/96 (0%)
M127	<i>Cpl29</i>	IDT double deletion	61/96 (64%)	79/96 (82%)	91/96 (95%)	1/96 (1%)
	<i>Crt1</i>		20/96 (22%)	12/96 (13%)	28/96 (29%)	
M127	<i>Cpl29</i>	Traditional double deletion	54/96 (56%)	20/96 (21%)	7/96 (7%)	0/96 (0%)
	<i>Crt1</i>		30/96 (31%)	0/96 (0%)	6/96 (6%)	
M123	<i>Cpl29</i>	IDT	72/96 (75%)	80/96 (83%)	94/96 (98%)	63/96 (66%)
M123	<i>Cpl29</i>	Traditional	45/96 (47%)	50/96 (52%)	48/96 (50%)	10/96 (10%)
M123	<i>Crt1</i>	IDT	42/96 (44%)	52/96 (54%)	43/96 (45%)	28/96 (29%)
M123	<i>Crt1</i>	Traditional	33/60 (55%)	11/60 (18%)	9/60 (15%)	2/60 (3%)
M123	<i>Cpl29</i>	IDT double deletion	65/96 (68%)	87/96 (91%)	85/96 (85%)	2/96 (2%)
	<i>Crt1</i>		45/96 (47%)	15/96 (16%)	16/96 (17%)	
M123	<i>Cpl29</i>	Traditional double deletion	37/87 (43%)	35/87 (40%)	30/87 (34%)	0/87 (0%)
	<i>Crt1</i>		22/87 (25%)	10/87 (11%)	6/87 (7%)	
M123	<i>Cpl29</i>	IDT, no gRNA ctrl	33/96 (34%)	61/96 (64%)	38/96 (40%)	0/96 (0%)
	<i>Crt1</i>		27/96 (28%)	13/96 (14%)	5/96 (5%)	

Table 15 Colony counts, number of clones, and transformation efficiencies of the IDT method in M127 and M123. Number of colonies on transformation plates calculated to total amount of positive colonies based on initial number of protoplasts and the likelihood of a positive clone as shown in Table 14.

Strain	Target	Method	# of colonies / plate	Total # of positive clones	Transformation efficiency (cfu/ μ g)
M127	<i>Cpl29</i>	IDT	130	$1,06 \times 10^4$	$1,32 \times 10^5$
M127	<i>Cpl29</i>	Traditional	100	$5,60 \times 10^1$	$1,40 \times 10^2$
M127	<i>Cpl29</i>	IDT, no gRNA ctrl	91	$2,55 \times 10^1$	$6,38 \times 10^1$
M127	<i>Crt1</i>	IDT	37 *	$1,04 \times 10^2$	$1,30 \times 10^3$
M127	<i>Crt1</i>	Traditional	337	0	0
M127	<i>Cpl29</i> & <i>Crt1</i>	IDT double deletion	112	$1,57 \times 10^2$	$1,96 \times 10^3$
M127	<i>Cpl29</i>	IDT, no gRNA ctrl	91	$2,55 \times 10^1$	$6,38 \times 10^1$
M123	<i>Cpl29</i>	IDT	505	$9,33 \times 10^3$ **	$2,34 \times 10^4$
M123	<i>Cpl29</i>	Traditional	312	$8,74 \times 10^2$	$2,19 \times 10^3$
M123	<i>Crt1</i>	IDT	138	$1,12 \times 10^3$	$2,81 \times 10^3$
M123	<i>Crt1</i>	Traditional	226	$1,90 \times 10^2$	$4,75 \times 10^2$
M123	<i>Cpl29</i> & <i>Crt1</i>	IDT double deletion	541	$3,03 \times 10^2$ **	$7,59 \times 10^2$
M123	<i>Cpl29</i> & <i>Crt1</i>	Traditional double deletion	139	0	0
M123	<i>Cpl29</i> & <i>Crt1</i>	IDT, no gRNA ctrl	234	0	0

*Average colony number based on six transformation plates due to high variability

**plates were full of small colonies, two quarters of three plates were counted and full plate colony number calculated based on the quarters.

From Table 14 we can clearly see that the IDT CRISPR/Cas9 method resulted in a significant increase in correctly integrated deletion cassettes. We can also observe that $\Delta cpl29$ seems to be much easier to introduce than $\Delta crt1$. This can simply be due to chromosomal arrangements preventing probably both the Cas9 and the dDNA from entering the *crt1*

locus. Also, an intact *crt1* may be crucial for the survival and growth of *T. reesei* and deletion strains are lost.

Based on Table 15 we can conclude that the CRISPR/Cas9 increased the transformation efficiency, depending on the deletion and the strain. The efficiency of $\Delta cpl29$ and $\Delta crt1$ in M127 increases about 1 000-fold with CRISPR/Cas9 in comparison to the non-CRISPR/Cas9 controls. In M123 the increase is not as dramatic, but still a notable 10 × and 6 × increase for $\Delta cpl29$ and $\Delta crt1$, respectively. For double deletions, a 2 000 × and 750 × increase in efficiency is observed for M127 and M123, respectively. The increases in efficiencies may be misleading and higher than in reality in cases where no positive clones were found in the control reactions. More clones should be screened for higher accuracy. This was, however, not possible in the time-frame of this study.

Not many conclusions can be made from the number of colonies per plate in Table 15. In some cases the total number of colonies is similar regardless of the method ($\Delta cpl29$ in M127, and $\Delta cpl29\Delta crt1$ in M127), in some cases the number of colonies was significantly decreased with the IDT method ($\Delta crt1$ in M127, $\Delta crt1$ in M123), and in some there was a clear increase in colony numbers with the IDT method ($\Delta cpl29$ in M123, $\Delta cpl29\Delta crt1$ in M123). However, the decrease in colony numbers is only noted in $\Delta crt1$ in both strains. This supports our idea that *crt1* is somewhat essential to the cells and the deletion is putting the cells at a disadvantage. We hypothesize that the $\Delta crt1$ is efficiently introduced by the IDT method, causing hindrance to the cell growth and/or survival. When the deletion was introduced without CRISPR/Cas9 a high number of background colonies, that still have at least one intact copy of *crt1*, grow on the plates.

If we look at the transformations in which the colony numbers stayed the same or increased, the differentiating factor was the strain in which the deletions were introduced. The amount of colonies stays roughly the same in M127 and increases in M123 when using the IDT method. This may be due to M127's inability to efficiently repair DSBs, and some cells are lost due to either unspecific DSBs or inefficient dDNA introduction. When assessing the colony numbers, it should be kept in mind that the calculations are only directive. Figure 19 presents images of the transformation plates and we can see that even though the total number of colonies is similar (e.g. $\Delta cpl29$ M127), the number of colonies that grow through

the plate are completely different. This may also be due to a difference in growth speed, but also shows that CRISPR/Cas9 provides advancement to the transformation and the successive screening, making it faster to achieve a final strain.

All in all, the clear increase in transformation efficiencies showed that the CRISPR/Cas9 system increased the number of transformed protoplasts, reducing the amount of background colonies and simplifying the screening for correct clones.

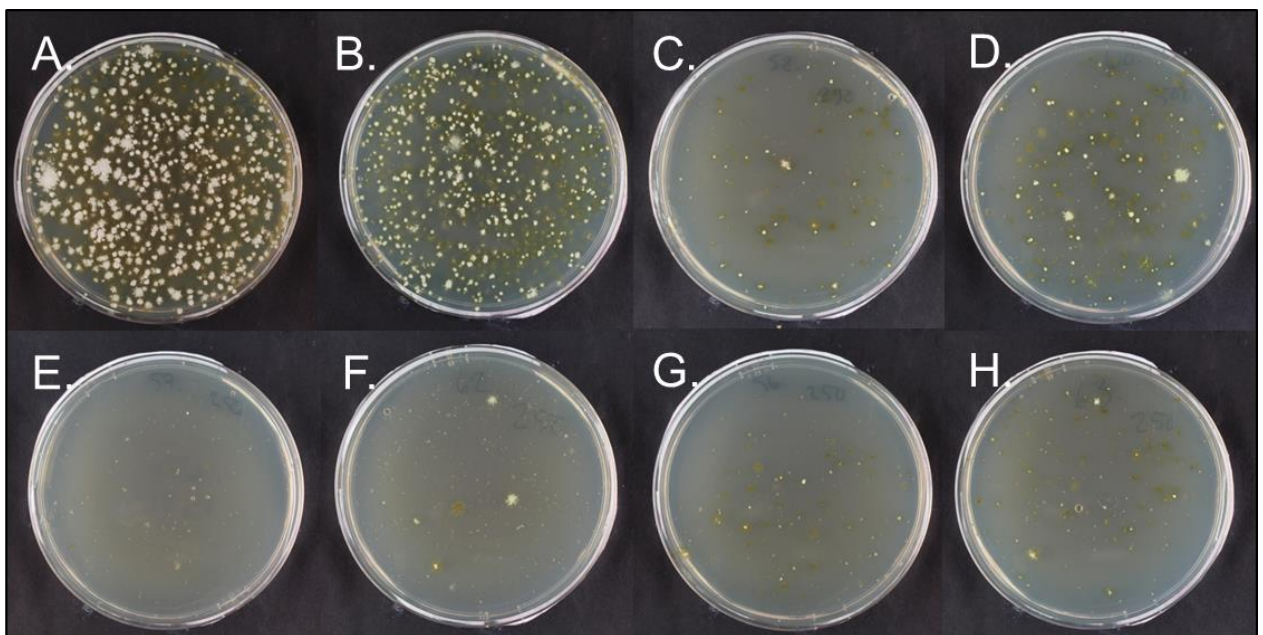


Figure 19 Transformation plates, 11 days post-transformation. **A.** IDT $\Delta cpl29$ M127; **B.** IDT $\Delta cpl29$ M123; **C.** IDT $\Delta crt1$ M127; **D.** IDT $\Delta crt1$ M123; **E.** Traditional $\Delta cpl29$ M127; **F.** Traditional $\Delta cpl29$ M123; **G.** Traditional $\Delta crt1$ M127; **H.** Traditional $\Delta crt1$ M123.

Based on the results, it seems that the low number of transformants in the no-gRNA control transformations of M127 is indeed due to the lack of *mus53*. It's important to keep in mind that low colony numbers of $\Delta cpl29$ in M127 acquired through the traditional method have not been noted previously, and is likely due to an error in the transformation method. The fact that M123 protoplasts recover better from the no-gRNA control reaction than M127 would suggest that the Cas9 protein alone introduces DSBs at random sites of the genome, which M127 is not able to repair. In the no-gRNA control transformation of M123 the number of transformants is similar to other clones; however with lower integration signals,

also suggesting random cut sites for Cas9. In this case the dDNA cassette could be incorrectly integrated at the random cut site and the *pyr4* marker would allow for the clones to survive.

To the best of our knowledge, no such effects have been published in other organisms and it could mean that *T. reesei* is somehow more sensitive to Cas9 than e.g. *P. chrysogenum*, in which no significant toxic effects were noted in ≤ 1520 nM Cas9 transformation concentrations (Y. Nygård, personal communication). If this is the case, it could also explain why *cas9* is difficult to express in *T. reesei*. There's a possibility that there are endogenous RNA's in *T. reesei* that can associate with Cas9 and introduce DSBs in unexpected locations. No similarity between the tracrRNA tail used in the sgRNAs of this study and *T. reesei* whole genome was found. Unfortunately the tracrRNA sequence of the IDT system is not public, so a nucleotide similarity comparison was not possible.

What's notable in Table 14 is that the method of identifying positive clones by strictly counting in only clones where all PCRs are positive can be misleading. There's always a possibility that a single PCR hasn't worked, and gives a false negative result. In almost all CRISPR/Cas9 transformations, the 5' integration signal gave the lowest success rate. This could be due to problems with the PCR or it could indicate that the 5' integration takes place less efficiently or in an unspecific manner. It can be that the target site close to the ORF 5' end isn't as effective as the target site in the 3' end. Also, in the case of *T. reesei*, some protoplasts contain multiple nuclei. This means that a positive ORF signal may be coming from just one nucleus, and the gene may be fully deleted in others. With proper purification steps, such a strain could become a clean deletion strain. Also, the very high number of ORF⁻ signals from the CRISPR/Cas9 deletions (especially in M127) suggests that the cassette is at least partly in the correct locus, suggesting that the CRISPR/Cas9 has induced a DSB in the locus as expected. With these possible analysis defects in mind, it's actually possible that the CRISPR/Cas9 is working even better than what this data leads us to believe. In *Δcpl29Δcrt1* M127, if one non-correct PCR is allowed, the number of correct clones increases from 1 to 8, while the number of correct double deletion clones from the traditional transformation remains 0. If one unsuccessful PCR is allowed for *Δcpl29Δcrt1* M123, the number of correct clones would be 8/96 (now 7/96) for the IDT method and 1/96 (now 0/96) for the traditional method. A more thorough analysis of the strains could be done through further purification steps.

Furthermore, it should be noted that no off-target effects have been analysed in this study. It's possible that the deletion cassette was incorrectly integrated in some cases, for example right next to the target site. This could result in a positive integration signal. The strains have not been studied further to see if there are multiple deletion cassette integrations at the target site or elsewhere in the genome. Especially in the case of the double deletions, both deletion cassettes included an identical *pyr4*-marker, which may have caused unwanted homologous recombination between the dDNAs themselves.

The double deletion does not work very efficiently. However, none of the traditional double deletions resulted in any positive clones for both deletions. In this sense, it may be plausible to use CRISPR/Cas9 to introduce two deletions at once, since screening 96 clones seems to provide potential double deleted strains. This would cut the time taken to delete multiple genes to half. It also may be possible to get higher double-deletion frequencies when using different target genes. From the CRISPR/Cas9 results we can see that *crt1* is quite difficult to delete, especially in M127. If both targets were as easily deleted as *cpl29*, a double deletion would maybe be obtained at a higher frequency, and perhaps even a triple deletion could be achieved in a single transformation.

5.7 Transporter Deletions

Putative sugar transporters *cpl27*, *cpl28* and *cpl29* and a lactose transporter *crt1* were deleted from M127 (*pyr4*⁻) with the traditional transformation method. *Cpl29* and *crt1* were also deleted using the CRISPR/Cas9 system. A selection of deletion transformants was grown in liquid culture to assess their cellulase and endoglucanase activities. It was of interest to know if any of these transporter deletions had any effect upon the secretion of cellulase enzymes, which would indicate that they were important in the sugar sensing pathways related to coordinating cellulase secretion.

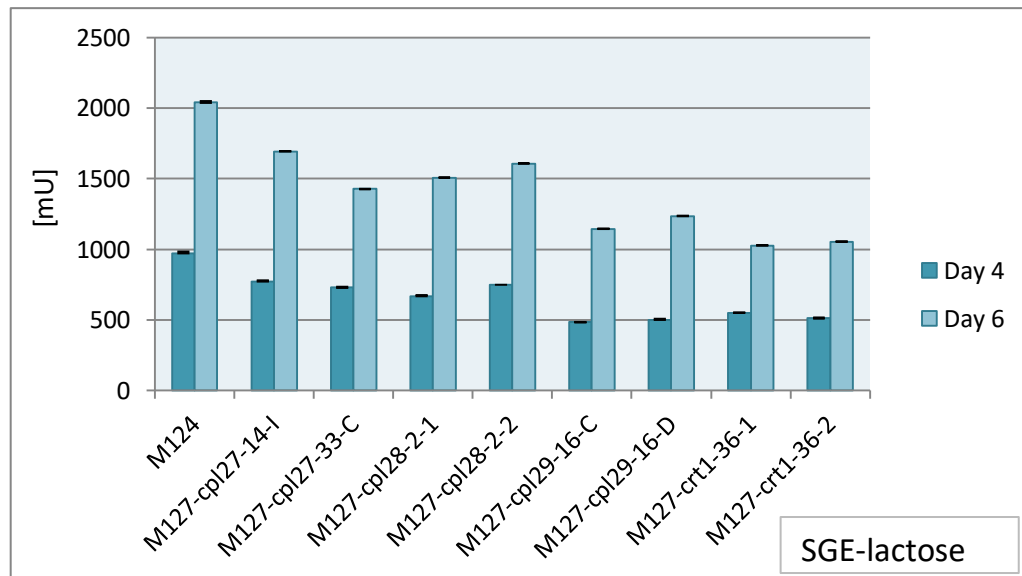


Figure 20 Azo-barley glucan assay for transporter deletions in SGE-lactose media. 1:500 sample dilution.

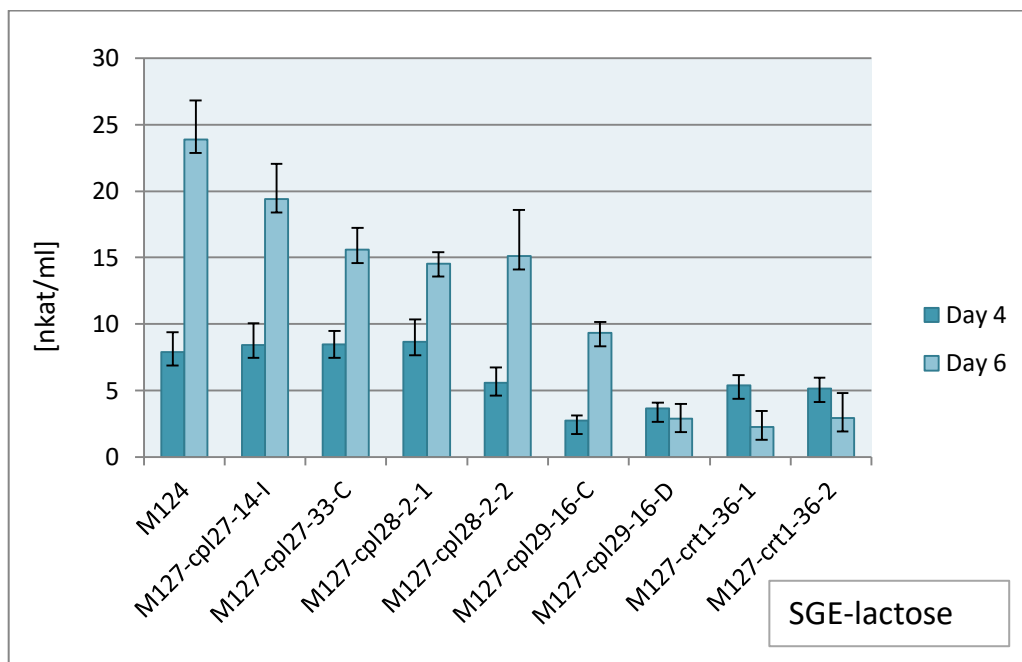


Figure 21 MULac assay for transporter deletions in SGE-lactose media. 1:100 (Day 4) & 1:500 (Day5) dilution.

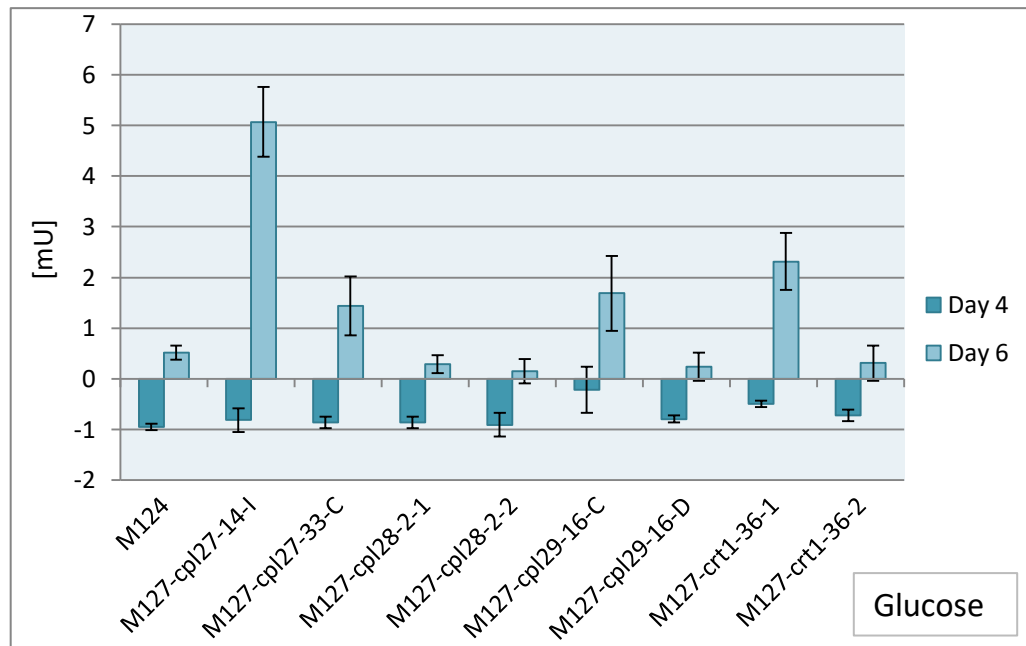


Figure 22 Azo-barley glucan assay for transporter deletions in glucose media. Samples were not diluted.

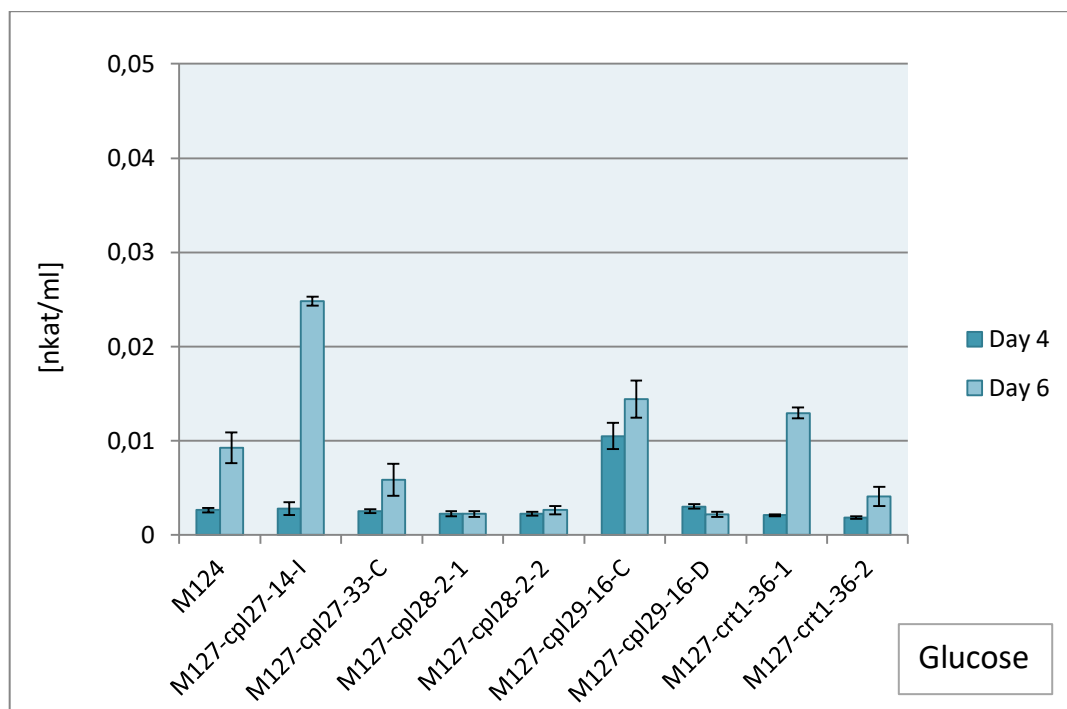


Figure 23 MULac assay for transporter deletions in glucose media. 1:2 dilution.

Figures 20 and 21 show the results of the Azo-barley glucan assay and MULac assay of the deletions strains in SGE-lactose media. Both assays show a similar pattern of activity, although Azo-barley glucan measures endoglucanases and MULac β -glucosidases and other enzymes. On day 4, the enzyme activities are quite similar between the control strain M124 and the deletions. A decrease in activity is only observed in $\Delta cpl29$ and $\Delta crt1$. Notably, all deletions show a decrease in cellulase production and/or secretion on day 6 when compared to the control strain M124.

Figures 22 and 23 show the results of the Azo-barley glucan assay and MULac assay of the deletion strains in glucose media. The enzyme activities in glucose are fairly similar between the control strain M124 and all deletion strains on day 4. It should be noted, however, that on day 4 the endoglucanase concentrations are so low that the Azo-barley glucan assay isn't exact enough to measure these concentrations, but instead recognizes them as negative values. However, this is an indication that on day 4 all samples are at a baseline of practically [0 mU] of endoglucanases. On day 6 some of the deletion strains' endoglucanase activity in glucose shoots up. From this data, we can see that a single $\Delta cpl27$ transformant shows increased endoglucanase activity in glucose media ($\Delta cpl27-14-I$). This same transformant shows a slightly increased activity in SGE-lactose media as well in comparison to its counterpart $\Delta cpl27-33-C$. Neither of these effects is however observed in the other $\Delta cpl27$ transformant ($\Delta cpl27-33-C$), and could simply be due to a measurement error. There is also a possibility that an additional mutation has taken place in one of the transformants and they aren't fully comparable. A similar effect can be seen in $\Delta cpl29$ and $\Delta crt1$ as well. Again, the activity of one of the transformants ($\Delta cpl29-16-C$ and $\Delta crt1-36-1$) is slightly higher in glucose when compared to its counterpart ($\Delta cpl29-16-D$ and $\Delta crt1-36-2$) deletions. The two clones seem to act fairly similarly in SGE-lactose.

It should be noted that all deletions are studied in two separate transformants, and consistently the increased activity in glucose is not recorded from all transformants. These deletion transformants should, however, be duplicates of one another. This calls for further investigation of the transformants. Are these changes in enzyme secretion actually due to the transporter deletions or some other mutation that has taken place or is it just a product of the assays' poor sensitivity in glucose media? With the MULac assay, $\Delta cpl29-16-C$ also shows slightly increased activity in SGE-lactose media in comparison to its counterpart

Δcpl29-16-D, which is not observed with the Azo-barley glucan assay. The strains should be checked with a Southern blot to specify the deletion and more replicates should be studied with enzyme assays. The two clones that seem to act consistently with each other are the *Δcpl28* clones. They are showing especially low activities in glucose.

In the case of *Δcrt1* and *Δcpl29* the seemingly increased endoglucanase activity in glucose media is likely due to the poor resolution of the assay in glucose media. The enzyme concentrations are so low that the assays may not be very reliable, and the observed changes in enzyme activity can be reduced to a measurement and statistical error. The fact that the *cpl27* deletion (*Δcpl27-14-I*) is showing an altered enzyme activity in both SGE-lactose and glucose medias, regardless of the assay in use, suggests that the effect may not only be due to a similar error of the method in use.

In any case, these mutations seem to introduce a lagging effect on the enzyme activities in SGE-lactose media. *Δcpl28* is disabling the enzyme production or secretion pathway in glucose media as well, while others may be introducing a slight increase of cellulase expression and/or secretion. The effect of *Δcpl28* is not as clear in the Azo-barley glucan assay, which may be due to the assay's poor performance in measuring very small amounts of endoglucanases, or *Δcpl28* may play a bigger role in regulating β -glucosidase expression and/or secretion than that of endoglucanases. Especially *Δcpl29* and *Δcrt1* on the other hand seem to lag the expression/secretion pathway in SGE-lactose and the effect is again clearer in the MULac results than the Azo-barley glucan. This could again be due to the deletions having specific effects on different cellulases, or simply the lower resolution of the Azo-barley glucan assay. For *Δcrt1* the low expression level on SGE-lactose is in line with currently published literature (Zhang, Kou et al. 2013, Ivanova, Baath et al. 2013).

For all measurements in glucose media, it should be noted that the observed changes are not very significant from a cellulase production view-point. Even for the increased expression of *Δcpl27-14-I*, the enzyme concentrations remain extremely low (5,1 mU and 0,24 nkat/ml).

Since no clear evidence of these putative transporters' role in glucose sensing can be drawn, the strains should be studied in regards to other carbon sources as well. As being said, the expression and/or secretion was hindered in SGE-lactose. Lactose typically induces the

cellulase production and *crt1* has previously been identified as a lactose transporter by Ivanova et al. (Ivanova, Baath et al. 2013). The low cellulase expression in $\Delta cp129$ in SGE-lactose media suggests that *cp129* may have a similar role and function as a lactose (or lactose hydrolysis product, galactose) transporter or sensor rather than glucose.

Unfortunately, there was no time to assess the deletions' effects on the strains' growth. One explanation to the different cellulase activities could be an altered growth speed in different media. *Cp128* for example has been recognized as a glucose transporter that does not have an effect on cellulase induction (Zhang, Kou et al. 2013). The observed low levels of cellulases in $\Delta cp128$ transformants could be due to hindered growth and a smaller number of cells.

6 Discussion

6.1 CRISPR/Cas9 in *T. reesei*

The results from this study suggest, but cannot unambiguously conclude that the Cas9 protein has significant toxic effects in *T. reesei*. The fact that the colony numbers are decreased when protoplasts are transformed with only the Cas9 protein and a dDNA fragment does indicate that Cas9 on its own is somehow not tolerated in *T. reesei*. The low number of colonies and lack of correct transformants in the AMA1 experiments support this as well. The further experiment comparing the strains M123 and M127 does suggest that the toxic effect is caused by the lack of *mus53* gene, involved in the NHEJ machinery of *T. reesei*. All of our efforts to obtain continuous *cas9* expression in *T. reesei* were done in M127. Similar experiments could be tested in M123 to find if the lack of *mus53* and the cells' inability to repair DSBs causes the problems in expression. In mammals the deletion of NHEJ machinery is a common practice to increase the likelihood of HR, also in CRISPR/Cas9 gene editing (Chu, Weber et al. 2015). To the best of our knowledge, no studies indicating notable cell death have been published. We cannot conclude if there are Cas9 associating endogenous RNAs in *T. reesei* as the IDT tracrRNA sequence is not public. If however there are no sequence similarities, the cell deaths suggest that Cas9 could be introducing cuts without a specific guide RNA or it may use other endogenous RNAs. Based on current literature this shouldn't be the case, but that Cas9 rather attaches and detaches to and from unspecific sites without inducing a DSB (Sternberg, Redding et al. 2014) (see Section 2.4.5, Off-target Effects). The high amount of transformants with the IDT method would, however, be in line with published literature suggesting that the efficiency to create an on-target DSB is much higher than the off-target DSBs (O'Geen, Henry et al. 2015).

To get answers to these questions, the off-target effects in *T. reesei* should be studied. If random cuts are introduced in high quantity in the absence of a gRNA, mutations should be created in both M123 and M127. The mutations could maybe be identified through ChIP-seq. A careful study of off-target effects could also shed light on continuous *cas9* expression systems. For gene editing purposes a continuous expression system isn't necessary since the

protein transformation seems to provide highly efficient gene editing results. A continuous expression is however crucial for dCas9 applications. It should be kept in mind that the dCas9 should under no circumstances introduce unspecific DSBs. In this sense, protein transformations of dCas9 would also be interesting to see if toxic effect occurs. That would suggest that something other than the endonuclease activity is causing the toxicity of Cas9 in *T. reesei* or that a transformation reagent is toxic. It's possible that instead of introducing unspecific DSBs, Cas9 is interfering with the intracellular RNAi machinery by simply binding RNAs (O'Connell, Oakes et al. 2014). Currently there are no commercial sources for the dCas9 so we were not able to investigate this theory. We have not been able to find any publications about using dCas9 in fungal species, although some publications using continuous *cas9* expression for gene editing in at least *T. reesei*, *P. chrysogenum*, and *A. niger* have been published (Pohl, Kiel et al. 2016, Nødvig, Nielsen et al. 2015, Liu, Chen et al. 2015).

Establishing a stable *cas9* expression may be much easier said than done. Weninger et al. (2016) tested 95 different combinations of gRNA expression systems, gRNA sequences, *cas9* expression promoters and *cas9* codon optimizations in *Pichia pastoris* and found only 6 combinations in which CRISPR/Cas9 mediated genome editing was efficient (Weninger, Hatzl et al. 2016). Of these 95 combinations, the functional *cas9* was human codon optimized. This reveals that finding the optimal conditions for expression can be very difficult, and at times illogical. Our study was unsuccessful in finding these optimal conditions for expressing CRISPR/Cas9 *in vivo* in *T. reesei*. Other promoters for *cas9* expression could be tested as well as other codon optimizations. The results of this study cannot demonstrate whether the promoters, codon optimizations or genome integration sites used were optimal or not. Based on the results from the *gpdA* promoter construct and the protein transformation, we suspect that *T. reesei* may be surprisingly intolerant towards Cas9. In the publication that used continuous *cas9* expression in *T. reesei*, the expression cassette may have been randomly integrated to a loci allowing for a very low level of expression and escaping rapid degradation or cell death (Liu, Chen et al. 2015). Notably the published study was done in the QM6a wild type isolate. Nevertheless, all studies so far suggest that once the CRISPR/Cas9 system is established and functioning, the genome editing is very efficient when compared to other methods.

Our results also indicate that the CRISPR/Cas9 as a gene editing method works very well in *T. reesei*. The fact that the protein transformation method seems to work very well in inducing deletions does offer some relief in new strain generation. Not only can the number of transformants to be screened be cut down, reducing the time and costs for strain generation, multiple deletions/insertions can possibly be introduced at once. In the future, using shorter flanks to the site of integration could be tested as well. Pohl et al. were able to initiate directed HR with only 60 bp flanks to the site of interest (Pohl, Kiel et al. 2016). If such short flanks could also be used in *T. reesei* the cloning steps to create dDNAs could be eliminated and dDNAs could be created in a single PCR reaction with primers including the flanks. The IDT method can also very easily be used in each individual transformation in any strain since there's no need to first set up a functional *cas9* and/or gRNA expression to make use of CRISPR/Cas9. This also relieves some pressure from the Cas9 toxicity and off-target effects, as the protein is not continuously present in the cells. According to literature, the use of protein transformation, as opposed to expression systems, lessens the off-target effects (Zuris, Thompson et al. 2015, Liang, Potter et al. 2015).

In the timeframe of this study there was no time to test more than two target genes and only deletions were tested. The lack of information about the chromosomal arrangement of *T. reesei* makes it difficult to assess why one of the chosen target genes was easier to abolish than the other. The published CRISPR/Cas9 efficiency studies mostly revolve around target site design, not target gene selection. In our case the problem probably doesn't lie in the gRNA target site, since obtaining a $\Delta crt1$ deletions strain was difficult with and without CRISPR/Cas9, but the efficiency was higher with CRISPR/Cas9.

What should be done next with the CRISPR/Cas9 development in *T. reesei* is to test this method for gene insertions as well. In this study, only deletions have been done. Also, the strains should be more carefully checked for off-target effects and multiple dDNA cassette integrations. In the case of double deletions (or insertions), a more careful look into the target sites and possible recombination between the dDNAs should be taken. Conclusions about using the same marker sequence in multiple dDNAs in one transformation cannot be drawn from the results presented above. However, if using the same marker in both dDNAs turns out to be problematic, different markers could be used. This would then slow down the marker recycling process, but would still potentially cut down the time needed for

creating strains with multiple modifications. However, now that we know that single deletions can be achieved in a very efficient manner in *T. reesei*, the option of using non-marker dDNAs could be revisited. This may not be an option for multiple simultaneous deletions, as the efficiency seems low based on current results, but the use of non-marker dDNAs would nevertheless reduce the time of new strain generation through sequential single gene modifications.

Introducing a deletion with the traditional transformation method typically takes three months. This includes the time it takes to transform the strain, purify a clean clone, and recycle the marker. Only after marker recycling can another modification be done. Protoplast preparation takes another week. In this sense, introducing a double deletion would reduce the time of strain creation to half. If the purpose of the deletion is to study the double deletions effects, no marker recycling would be necessary either.

Another approach to find exactly how much CRISPR/Cas9 can reduce the time to create new strains is to purify the correct clones. The PCR products were very clean, and it may be that the CRISPR/Cas9 produced transformants are so clean that only one purification step could sufficiently produce a clean strain. This would cut the time needed for strain purification in half.

6.2 Transporter Deletions

As *T. reesei* is widely used as a platform for both cellulase and heterologous protein production, the importance of understanding the agents involved in protein production capacities are tremendous. Cellulase expression is a self-closing loop, and shuts down once enough cellulases are present to hydrolyse cellulose into glucose to allow for sufficient cell growth. Because the cellulase expression system is so efficient, the promoters involved in this pathway are often used to produce heterologous proteins as well (Schuster, Schmoll 2010). In this sense more thorough understanding of the regulation of these promoters holds great promise when using *T. reesei* as a production host. This study focused on putative glucose transporters as an attempt to shut down the cells' glucose sensing capabilities and freeing the cellulase expression system from glucose repression. This was

not successfully done, and instead the studied putative transporters may play a role in sensing and transporting other carbon sources, or they may be involved in something completely different. *Cpl27* and *cpl29* were not previously known to be involved in glucose or sugar sensing. They may simply be membrane transporters for sugars. As this study mainly focused on successful genome editing of *T. reesei* using CRISPR/Cas9, the transporter deletion strains weren't studied in extensive detail. Based on the results, $\Delta cpl28$ is involved in glucose transporting or sensing. The absence of this gene seemed to shut the cellulase expression and/or secretion machinery down, but as suggested by Zhang et al. (2013), *cpl28* could be independent of the cellulase expression system and the observed effect could be due to measurement error or hindered growth (Zhang, Kou et al. 2013). It could be interesting to see what how a strain with *cpl28* overexpressed would act in glucose media.

7 Conclusions

To our best knowledge, this is the first published study showing a successful CRISPR/Cas9 gene deletion in *T. reesei* using a fully synthetic, transient transformation method. This new approach of introducing deletions shows great promise in high transformation efficiencies, lessening the time spent on creating a new deletion strain. There is no indication of why this method couldn't be used for knock-in modifications as well. This is also the first time that two deletions have ever been done simultaneously.

Out of four different sugar transporter/sensor gene deletions, none were found to significantly benefit the cellulase expression and secretion system. Unfortunately due to time constraints, these genes' functions were not identified and elucidated in detail.

8 Acknowledgements

I would like to thank the Technical Research Centre of Finland (VTT) for the thesis opportunity. Special thanks to my thesis instructors team leader Christopher Landowski and senior scientist Mari Valkonen, and the group supervisor Markku Saloheimo for invaluable help in performing the experiments and writing the finished thesis. I would also like to thank the university lecturer Suvi Taira from the University of Helsinki for her help with the writing process.

Thank you to Yvonne Nygård from the University of Groningen for providing insight to the CRISPR/Cas9 system in filamentous fungi. Thank you to all the researchers who took part to project meetings and gave invaluable comments and ideas for the thesis.

Also, thank you to Rui Liu, Ling Chen, Yanping Jiang, Zhihua Zhou, and Gen Zou from CAS-Key Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, China for providing a *T. reesei* optimized *cas9* expression plasmid. Thank you to Christina Nødvig, Jakob Nielsen, Martin Kogle, and Uffe Mortensen from the Technical University of Denmark for providing the *cas9* expressing AMA1 plasmid.

This thesis was funded by the TEKES Living Factories programme (<http://www.vtt.fi/sites/livingfactories/en>). I would like to thank the project coordinator research professor Merja Penttilä from VTT for this opportunity.

9 References

- CRISPR Cas9 - gRNA Design. 2016. **Accessed June 6th, 2016**(https://www.abmgood.com/marketing/knowledge_base/CRISPR_Cas9_gRNA_Design.php#2),.
- CRISPR-Cas9 Genome Editing. 2016. **Accessed July 19th, 2016**(<https://eu.idtdna.com/pages/products/genome-editing/crispr-cas9>),.
- ARO, N., PAKULA, T. and PENTTILA, M., 2005. Transcriptional regulation of plant cell wall degradation by filamentous fungi. *FEMS microbiology reviews*, **29**(4), pp. 719-739.
- BAE, S., PARK, J. and KIM, J.S., 2014. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics (Oxford, England)*, **30**(10), pp. 1473-1475.
- BEERLI, R.R. and BARBAS, C.F., 2002. Engineering polydactyl zinc-finger transcription factors. *Nat Biotech*, **20**(2), pp. 135-141.
- BIBIKOVA, M., BEUMER, K., TRAUTMAN, J.K. and CARROLL, D., 2003. Enhancing gene targeting with designed zinc finger nucleases. *Science (New York, N.Y.)*, **300**(5620), pp. 764.
- BIBIKOVA, M., CARROLL, D., SEGAL, D.J., TRAUTMAN, J.K., SMITH, J., KIM, Y.G. and CHANDRASEGARAN, S., 2001. Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Molecular and cellular biology*, **21**(1), pp. 289-297.
- BIBIKOVA, M., GOLIC, M., GOLIC, K.G. and CARROLL, D., 2002. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics*, **161**(3), pp. 1169-1175.
- BISCHOF, R.H., RAMONI, J. and SEIBOTH, B., 2016. Cellulases and beyond: the first 70Â years of the enzyme producer Trichoderma reesei. *Microbial Cell Factories*, **15**(1), pp. 1-13.
- BOCH, J., SCHOLZE, H., SCHORNACK, S., LANDGRAF, A., HAHN, S., KAY, S., LAHAYE, T., NICKSTADT, A. and BONAS, U., 2009. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science*, **326**(5959), pp. 1509-1512.
- BONAS, U., STALL, R.E. and STASKAWICZ, B., 1989. Genetic and structural characterization of the avirulence gene avrBs3 from Xanthomonas campestris pv. vesicatoria. *Molecular and General Genetics MGG*, **218**(1), pp. 127-136.
- BORTESI, L. and FISCHER, R., 2015. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, **33**(1), pp. 41-52.
- CAPECCHI, M.R., 1989. Altering the genome by homologous recombination. *Science (New York, N.Y.)*, **244**(4910), pp. 1288-1292.

- CATHOMEN, T. and KEITH JOUNG, J., 2008. Zinc-finger Nucleases: The Next Generation Emerges. *Molecular therapy : the journal of the American Society of Gene Therapy*, **16**(7), pp. 1200-1207.
- CENCIC, R., MIURA, H., MALINA, A., ROBERT, F., ETHIER, S., SCHMEING, T.M., DOSTIE, J. and PELLETIER, J., 2014. Protospacer Adjacent Motif (PAM)-Distal Sequences Engage CRISPR Cas9 DNA Target Cleavage. *PLoS ONE*, **9**(10), pp. e109213.
- CERMAK, T., DOYLE, E.L., CHRISTIAN, M., WANG, L., ZHANG, Y., SCHMIDT, C., BALLER, J.A., SOMIA, N.V., BOGDANOVE, A.J. and VOYTAS, D.F., 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids research*, **39**(12), pp. e82-e82.
- CHEN, B., GILBERT, L., CIMINI, B., SCHNITZBAUER, J., ZHANG, W., LI, G., PARK, J., BLACKBURN, E., WEISSMAN, J., QI, L. and HUANG, B., 2013. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell*, **155**(7), pp. 1479-1491.
- CHERRY, J.R. and FIDANTSEF, A.L., 2003. Directed evolution of industrial enzymes: an update. *Curr Opin Biotechnol*, **14**.
- CHO, S.W., KIM, S., KIM, Y., KWEON, J., KIM, H.S., BAE, S. and KIM, J.S., 2014. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome research*, **24**(1), pp. 132-141.
- CHRISTIAN, M., CERMAK, T., DOYLE, E.L., SCHMIDT, C., ZHANG, F., HUMMEL, A., BOGDANOVE, A.J. and VOYTAS, D.F., 2010. Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics*, **186**(2), pp. 757-761.
- CHU, V.T., WEBER, T., WEFERS, B., WURST, W., SANDER, S., RAJEWSKY, K. and KUHN, R., 2015. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotech*, **33**(5), pp. 543-548.
- COHN, M. and HORIBATA, K., 1959. Physiology of the Inhibition by Glucose of the Induced Synthesis of the β -Galactoside-Enzyme System of Escherichia Coli. *Journal of Bacteriology*, **78**(5), pp. 624-635.
- CORTEZ, C., 2015. CRISPR 101- Homology-Directed Repair. **Accessed May 26th, 2016**(<http://blog.addgene.org/crispr-101-homology-directed-repair>),.
- CRADICK, T.J., FINE, E.J., ANTICO, C.J. and BAO, G., 2013. CRISPR/Cas9 systems targeting β -globin and CCR5 genes have substantial off-target activity. *Nucleic acids research*, **41**(20), pp. 9584-9592.
- DANG, Y., JIA, G., CHOI, J., MA, H., ANAYA, E., YE, C., SHANKAR, P. and WU, H., 2015. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome biology*, **16**(1), pp. 1-10.

- DASHTBAN, M., BUCHKOWSKI, R. and QIN, W., 2011. Effect of different carbon sources on cellulase production by *Hypocrea jecorina* (*Trichoderma reesei*) strains. *International Journal of Biochemistry and Molecular Biology*, **2**(3), pp. 274-286.
- DELGADO-JARANA, J., MORENO-MATEOS, M. and BENÍTEZ, T., 2003. Glucose Uptake in *Trichoderma harzianum*: Role of gtt1. *Eukaryotic Cell*, **2**(4), pp. 708-717.
- DICARLO, J.E., NORVILLE, J.E., MALI, P., RIOS, X., AACH, J. and CHURCH, G.M., 2013. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic acids research*, **41**(7), pp. 4336-4343.
- DOENCH, J., 2016. How to Design Your gRNA for CRISPR Genome Editing. **Accessed June 13th, 2016**(<http://blog.addgene.org/how-to-design-your-grna-for-crispr-genome-editing>),.
- DOS SANTOS CASTRO, L., DE PAULA, R.G., ANTONIÊTO, A.C.C., PERSINOTI, G.F., SILVA-ROCHA, R. and SILVA, R.N., 2016. Understanding the Role of the Master Regulator XYR1 in *Trichoderma reesei* by Global Transcriptional Analysis. *Frontiers in Microbiology*, **7**, pp. 10.3389/fmicb.2016.00175.
- DOUDNA, J.A. and CHARPENTIER, E., 2014. The new frontier of genome engineering with CRISPR-Cas9. *Science*, **346**(6213),.
- DREIER, B., SEGAL, D.J. and BARBAS, C.F., 3rd, 2000. Insights into the molecular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains. *Journal of Molecular Biology*, **303**(4), pp. 489-502.
- DUAN, J., LU, G., XIE, Z., LOU, M., LUO, J., GUO, L. and ZHANG, Y., 2014. Genome-wide identification of CRISPR/Cas9 off-targets in human genome. *Cell research*, **24**(8), pp. 1009-1012.
- EIGES, R., SCHULDINER, M., DRUKKER, M., YANUKA, O., ITSKOVITZ-ELDOR, J. and BENVENISTY, N., 2001. Establishment of human embryonic stem cell-transfected clones carrying a marker for undifferentiated cells. *Current Biology*, **11**(7), pp. 514-518.
- ELROD-ERICKSON, M., ROULD, M.A., NEKLUDOVA, L. and PABO, C.O., 1996. Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure (London, England : 1993)*, **4**(10), pp. 1171-1180.
- FARZADFARD, F., PERLI, S.D. and LU, T.K., 2013. Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS synthetic biology*, **2**(10), pp. 604-613.
- FU, Y., FODEN, J.A., KHAYTER, C., MAEDER, M.L., REYON, D., JOUNG, J.K. and SANDER, J.D., 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature biotechnology*, **31**(9), pp. 822-826.
- FU, Y., SANDER, J.D., REYON, D., CASCIO, V.M. and JOUNG, J.K., 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotech*, **32**(3), pp. 279-284.

GAJ, T., GERSBACH, C.A. and BARBAS, C.F., 2013. ZFN, TALEN and CRISPR/Cas-based methods for genome engineering. *Trends in biotechnology*, **31**(7), pp. 397-405.

GAO, Y. and ZHAO, Y., 2014. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *Journal of Integrative Plant Biology*, **56**(4), pp. 343-349.

GILBERT, L.A., HORLBECK, M.A., ADAMSON, B., VILLALTA, J.E., CHEN, Y. and WHITEHEAD, E.H., 2014. Genome-scale CRISPR-mediated control of gene repression and activation. *Cell*, **159**.

GILBERT, L., LARSON, M., MORSUT, L., LIU, Z., BRAR, G., TORRES, S., STERN-GINOSSAR, N., BRANDMAN, O., WHITEHEAD, E., DOUDNA, J., LIM, W., WEISSMAN, J. and QI, L., 2013. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*, **154**(2), pp. 442-451.

GUPTA, R. and MUSUNURU, K., 2014. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *The Journal of clinical investigation*, **124**(10), pp. 4154-4161.

HÄKKINEN, M., VALKONEN, M.J., WESTERHOLM-PARVINEN, A., ARO, N., ARVAS, M., VITIKAINEN, M., PENTTILÄ, M., SALOHEIMO, M. and PAKULA, T.M., 2014. Screening of candidate regulators for cellulase and hemicellulase production in *Trichoderma reesei* and identification of a factor essential for cellulase production. *Biotechnology for biofuels*, **7**(1), pp. 14-6834-7-14.

HERBERS, K., CONRAD-STAUCH, J. and BONAS, U., 1992. Race-specificity of plant resistance to bacterial spot disease determined by repetitive motifs in a bacterial avirulence protein. *Nature*, **356**(6365), pp. 172-174.

HILTON, I.B., D'IPPOLITO, A.M., VOCKLEY, C.M., THAKORE, P.I., CRAWFORD, G.E., REDDY, T.E. and GERSBACH, C.A., 2015. Epigenome editing by a CRISPR/Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature biotechnology*, **33**(5), pp. 510-517.

HIMMEL, M.E., DING, S.Y., JOHNSON, D.K., ADNEY, W.S., NIMLOS, M.R., BRADY, J.W. and FOUST, T.D., 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science (New York, N.Y.)*, **315**(5813), pp. 804-807.

HSU, P.D., SCOTT, D.A., WEINSTEIN, J.A., RAN, F.A., KONERMANN, S., AGARWALA, V., LI, Y., FINE, E.J., WU, X., SHALEM, O., CRADICK, T.J., MARRAFFINI, L.A., BAO, G. and ZHANG, F., 2013a. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotech*, **31**(9), pp. 827-832.

HSU, P.D., SCOTT, D.A., WEINSTEIN, J.A., RAN, F.A., KONERMANN, S., AGARWALA, V., LI, Y., FINE, E.J., WU, X., SHALEM, O., CRADICK, T.J., MARRAFFINI, L.A., BAO, G. and ZHANG, F., 2013b. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology*, **31**(9), pp. 827-832.

- ILMEN, M., ONNELA, M.L., KLEMSDAL, S., KERANEN, S. and PENTTILA, M., 1996. Functional analysis of the cellobiohydrolase I promoter of the filamentous fungus *Trichoderma reesei*. *Molecular & general genetics : MGG*, **253**(3), pp. 303-314.
- ILMÉN, M., SALOHEIMO, A., ONNELA, M.L. and PENTTILÄ, M.E., 1997. Regulation of cellulase gene expression in the filamentous fungus *Trichoderma reesei*. *Applied and Environmental Microbiology*, **63**(4), pp. 1298-1306.
- IVANOVA, C., BAATH, J.A., SEIBOTH, B. and KUBICEK, C.P., 2013. Systems analysis of lactose metabolism in *Trichoderma reesei* identifies a lactose permease that is essential for cellulase induction. *PloS one*, **8**(5), pp. e62631.
- JIANG, W., BRUEGGEMAN, A.J., HORKEN, K.M., PLUCINAK, T.M. and WEEKS, D.P., 2014. Successful Transient Expression of Cas9 and Single Guide RNA Genes in *Chlamydomonas reinhardtii*. *Eukaryotic Cell*, **13**(11), pp. 1465-1469.
- JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J.A. and CHARPENTIER, E., 2012. A Programmable Dual-RNA\textendashGuided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, **337**(6096), pp. 816-821.
- KIM, J.H., ROY, A., JOUANDOT, D. and CHO, K.H., 2013. The glucose signaling network in yeast. *Biochimica et biophysica acta*, **1830**(11), pp. 5204-5210.
- KONERMANN, S., BRIGHAM, M.D., TREVINO, A.E., JOUNG, J., ABUDAYYEH, O.O., BARCENA, C., HSU, P.D., HABIB, N., GOOTENBERG, J.S., NISHIMASU, H., NUREKI, O. and ZHANG, F., 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, **517**(7536), pp. 583-588.
- KRAPPMANN, S., 2007. Gene targeting in filamentous fungi: the benefits of impaired repair. *Fungal Biology Reviews*, **21**(1), pp. 25-29.
- KUBODERA, T., YAMASHITA, N. and NISHIMURA, A., 2002. Transformation of *Aspergillus* sp. and *Trichoderma reesei* using the pyrithiamine resistance gene (*ptrA*) of *Aspergillus oryzae*. *Bioscience, biotechnology, and biochemistry*, **66**(2), pp. 404-406.
- KUMAR, R., SINGH, S. and SINGH, O.V., 2008. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *Journal of Industrial Microbiology {\&}* *Biotechnology*, **35**(5), pp. 377-391.
- LANDER, E.S., LINTON, L.M., BIRREN, B., NUSBAUM, C., ZODY, M.C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., MELDRIM, J., MESIROV, J.P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, Y., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L.,

MERCER, S., MILNE, S., MULLIKIN, J.C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R.H., WILSON, R.K., HILLIER, L.W., MCPHERSON, J.D., MARRA, M.A., MARDIS, E.R., FULTON, L.A., CHINWALLA, A.T., PEPIN, K.H., GISH, W.R., CHISSOE, S.L., WENDL, M.C., DELEHAUNTY, K.D., MINER, T.L., DELEHAUNTY, A., KRAMER, J.B., COOK, L.L., FULTON, R.S., JOHNSON, D.L., MINX, P.J., CLIFTON, S.W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S., SLEZAK, T., DOGGETT, N., CHENG, J.F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., GIBBS, R.A., MUZNY, D.M., SCHERER, S.E., BOUCK, J.B., SODERGREN, E.J., WORLEY, K.C., RIVES, C.M., GORRELL, J.H., METZKER, M.L., NAYLOR, S.L., KUCHERLAPATI, R.S., NELSON, D.L., WEINSTOCK, G.M., SAKAKI, Y., FUJIYAMA, A., HATTORI, M., YADA, T., TOYODA, A., ITOH, T., KAWAGOE, C., WATANABE, H., TOTOKI, Y., TAYLOR, T., WEISSENBAACH, J., HEILIG, R., SAURIN, W., ARTIGUENAVE, F., BROTTIER, P., BRULS, T., PELLETIER, E., ROBERT, C., WINCKER, P., SMITH, D.R., DOUCETTE-STAMM, L., RUBENFIELD, M., WEINSTOCK, K., LEE, H.M., DUBOIS, J., ROSENTHAL, A., PLATZER, M., NYAKATURA, G., TAUDIEN, S., RUMP, A., YANG, H., YU, J., WANG, J., HUANG, G., GU, J., HOOD, L., ROWEN, L., MADAN, A., QIN, S., DAVIS, R.W., FEDERSPIEL, N.A., ABOLA, A.P., PROCTOR, M.J., MYERS, R.M., SCHMUTZ, J., DICKSON, M., GRIMWOOD, J., COX, D.R., OLSON, M.V., KAUL, R., RAYMOND, C., SHIMIZU, N., KAWASAKI, K., MINOSHIMA, S., EVANS, G.A., ATHANASIOU, M., SCHULTZ, R., ROE, B.A., CHEN, F., PAN, H., RAMSER, J., LEHRACH, H., REINHARDT, R., MCCOMBIE, W.R., DE LA BASTIDE, M., DEDHIA, N., BLOCKER, H., HORNISCHER, K., NORDSIEK, G., AGARWALA, R., ARAVIND, L., BAILEY, J.A., BATEMAN, A., BATZOGLOU, S., BIRNEY, E., BORK, P., BROWN, D.G., BURGE, C.B., CERUTTI, L., CHEN, H.C., CHURCH, D., CLAMP, M., COPLEY, R.R., DOERKS, T., EDDY, S.R., EICHLER, E.E., FUREY, T.S., GALAGAN, J., GILBERT, J.G., HARMON, C., HAYASHIZAKI, Y., HAUSSLER, D., HERMJAKOB, H., HOKAMP, K., JANG, W., JOHNSON, L.S., JONES, T.A., KASIF, S., KASPRYZK, A., KENNEDY, S., KENT, W.J., KITTS, P., KOONIN, E.V., KORF, I., KULP, D., LANCET, D., LOWE, T.M., MCLYSAGHT, A., MIKKELSEN, T., MORAN, J.V., MULDER, N., POLLARA, V.J., PONTING, C.P., SCHULER, G., SCHULTZ, J., SLATER, G., SMIT, A.F., STUPKA, E., SZUSTAKOWKI, J., THIERRY-MIEG, D., THIERRY-MIEG, J., WAGNER, L., WALLIS, J., WHEELER, R., WILLIAMS, A., WOLF, Y.I., WOLFE, K.H., YANG, S.P., YEH, R.F., COLLINS, F., GUYER, M.S., PETERSON, J., FELSENFELD, A., WETTERSTRAND, K.A., PATRINOS, A., MORGAN, M.J., DE JONG, P., CATANESE, J.J., OSOEGAWA, K., SHIZUYA, H., CHOI, S., CHEN, Y.J., SZUSTAKOWKI, J. and INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM, 2001. Initial sequencing and analysis of the human genome. *Nature*, **409**(6822), pp. 860-921.

LE CROM, S., SCHACKWITZ, W., PENNACCHIO, L., MAGNUSON, J.K., CULLEY, D.E., COLLETT, J.R., MARTIN, J., DRUZHININA, I.S., MATHIS, H., MONOT, F., SEIBOTH, B., CHERRY, B., REY, M., BERKA, R., KUBICEK, C.P., BAKER, S.E. and MARGEOT, A., 2009. Tracking the roots of cellulase hyperproduction by the fungus *Trichoderma reesei* using massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(38), pp. 16151-16156.

LIANG, X., POTTER, J., KUMAR, S., ZOU, Y., QUINTANILLA, R., SRIDHARAN, M., CARTE, J., CHEN, W., ROARK, N., RANGANATHAN, S., RAVINDER, N. and CHESNUT, J.D., 2015. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *Journal of Biotechnology*, **208**, pp. 44-53.

- LIN, Y., CRADICK, T.J., BROWN, M.T., DESHMUKH, H., RANJAN, P., SARODE, N., WILE, B.M., VERTINO, P.M., STEWART, F.J. and BAO, G., 2014. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic acids research*, .
- LIU, R., CHEN, L., JIANG, Y., ZHOU, Z. and ZOU, G., 2015. Efficient genome editing in filamentous fungus *Trichoderma reesei* using the CRISPR/Cas9 system.
- MA, M., YE, A.Y., ZHENG, W. and KONG, L., 2013. A Guide RNA Sequence Design Platform for the CRISPR/Cas9 System for Model Organism Genomes. *BioMed Research International*, **2013**, pp. 10.1155/2013/270805.
- MANDELS, M. and REESE, E.T., 1960. Induction of Cellulase in Fungi by Cellobiose. *Journal of Bacteriology*, **79**(6), pp. 816-826.
- MARRAFFINI, L., 2015. *CRISPR-Cas immunity in prokaryotes*. Nature: Nature.
- MARTINEZ, D., BERKA, R., HENRISSAT, B., SALOHEIMO, M., ARVAS, M., BAKER, S., CHAPMAN, J., CHERTKOV, O., COUTINHO, P., CULLEN, D., DANCHIN, E., GRIGORIEV, I., HARRIS, P., JACKSON, M., KUBICEK, C., HAN, C., HO, I., LARRONDO, L., DE LEON, A., MAGNUSON, J., MERINO, S., MISRA, M., NELSON, B., PUTNAM, N., ROBERTSE, B., SALAMOV, A., SCHMOLL, M., TERRY, A., THAYER, N., WESTERHOLM-PARVINEN, A., SCHOCH, C., YAO, J., BARABOTE, R., NELSON, M., DETTER, C., BRUCE, D., KUSKE, C., XIE, G., RICHARDSON, P., ROKHSAR, D., LUCAS, S., RUBIN, E., DUNN-COLEMAN, N., WARD, M. and BRETTIN, T., 2008. *Genome sequencing and analysis of the biomass-degrading fungus Trichoderma reesei (syn. Hypocrea jecorina)*. 26 edn. Nature Biotechnology: Nature Biotechnology.
- MASHIMO, T., 2014. Gene targeting technologies in rats: zinc finger nucleases, transcription activator-like effector nucleases, and clustered regularly interspaced short palindromic repeats. *Development, growth & differentiation*, **56**(1), pp. 46-52.
- MEYER, V., 2008. Genetic engineering of filamentous fungi--progress, obstacles and future trends. *Biotechnology Advances*, **26**(2), pp. 177-185.
- MOHR, S.E., HU, Y., EWEN-CAMPEN, B., HOUSDEN, B.E., VISWANATHA, R. and PERRIMON, N., 2016. CRISPR guide RNA design for research applications. *The FEBS journal*, .
- MOORE, R., CHANDRAHAS, A. and BLERIS, L., 2014. Transcription Activator-like Effectors: A Toolkit for Synthetic Biology. *ACS Synthetic Biology*, **3**(10), pp. 708-716.
- MORBITZER, R., RÖMER, P., BOCH, J. and LAHAYE, T., 2010. Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(50), pp. 21617-21622.
- MOSCOU, M.J. and BOGDANOVE, A.J., 2009. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science*, **326**(5959), pp. 1501-1501.

- MUKHERJEE, P.K., HORWITZ, B.A., SINGH, U.S., MUKHERJEE, M. and SCHMOLL, M., 2013. *Trichoderma: Biology and Applications*. CABI.
- NAITO, Y., HINO, K., BONO, H. and UI-TEI, K., 2015. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics*, **31**(7), pp. 1120-1123.
- NODVIG, C.S., NIELSEN, J.B., KOGLE, M.E. and MORTENSEN, U.H., 2015. A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. *PloS one*, **10**(7), pp. e0133085.
- NØDVIG, C.S., NIELSEN, J.B., KOGLE, M.E. and MORTENSEN, U.H., *A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi*. - Public Library of Science.
- O'CONNELL, M.R., OAKES, B.L., STERNBERG, S.H., EAST-SELETSKY, A., KAPLAN, M. and DOUDNA, J.A., 2014. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature*, **516**(7530), pp. 263-266.
- O'GEEN, H., HENRY, I.M., BHAKTA, M.S., MECKLER, J.F. and SEGAL, D.J., 2015. A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. *Nucleic acids research*, **43**(6), pp. 3389-3404.
- OUSTEROUT, D.G., KABADI, A.M., THAKORE, P.I., MAJOROS, W.H., REDDY, T.E. and GERSBACH, C.A., 2015. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nature communications*, **6**, pp. 6244.
- ÖZCAN, S. and JOHNSTON, M., 1999. Function and Regulation of Yeast Hexose Transporters. *Microbiology and Molecular Biology Reviews*, **63**(3), pp. 554-569.
- PALOHEIMO, M., HAARMANN, T., MÄKINEN, S. and VEHEMAANPERÄ, J., 2016. Production of Industrial Enzymes in *Trichoderma reesei*. In: C.D. MONIKA SCHMOLL, ed, *Gene Expression Systems in Fungi: Advancements and Applications*. Switzerland: Springer, pp. 23.
- PATTANAYAK, V., LIN, S., GUILINGER, J.P., MA, E., DOUDNA, J.A. and LIU, D.R., 2013. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotech*, **31**(9), pp. 839-843.
- PAVLETICH, N.P. and PABO, C.O., 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science (New York, N.Y.)*, **252**(5007), pp. 809-817.
- PENNISI, E., 2013. The CRISPR Craze. *Science*, **341**(6148), pp. 833-836.
- PENTTILÄ, M., NEVALAINEN, H., RATTO, M., SALMINEN, E. and KNOWLES, J., 1987. A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. *Gene*, **61**(2), pp. 155-164.
- PETERSON, R. and NEVALAINEN, H., 2012. *Trichoderma reesei* RUT-C30--thirty years of strain improvement. *Microbiology (Reading, England)*, **158**(Pt 1), pp. 58-68.

- PLATT, R., CHEN, S., ZHOU, Y., YIM, M., SWIECH, L., KEMPTON, H., DAHLMAN, J., PARNAS, O., EISENHAURE, T., JOVANOVIĆ, M., GRAHAM, D., JHUNJHUNWALA, S., HEIDENREICH, M., XAVIER, R., LANGER, R., ANDERSON, D., HACOEN, N., REGEV, A., FENG, G., SHARP, P. and ZHANG, F., 2014. CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. *Cell*, **159**(2), pp. 440-455.
- POHL, C., KIEL, J.A.K.W., DRIESSEN, A.J.M., BOVENBERG, R.A.L. and NYGÅRD, Y., 2016. CRISPR/Cas9 Based Genome Editing of *Penicillium chrysogenum*. *ACS Synthetic Biology*, **0**(0), pp. nu.
- PUCHTA, H., DUJON, B. and HOHN, B., 1993. Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic acids research*, **21**(22), pp. 5034-5040.
- QI, L.S., LARSON, M.H., GILBERT, L.A., DOUDNA, J.A., WEISSMAN, J.S., ARKIN, A.P. and LIM, W.A., 2013. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, **152**(5), pp. 1173-1183.
- RAMOS, A.S., CHAMBERGO, F.S., BONACCORSI, E.D., FERREIRA, A.J., CELLA, N., GOMBERT, A.K., TONSO, A. and EL-DORRY, H., 2006. Oxygen- and glucose-dependent expression of *Trhxt1*, a putative glucose transporter gene of *Trichoderma reesei*. *Biochemistry*, **45**(26), pp. 8184-8192.
- RAN, F.A., HSU, P.D., LIN, C., GOOTENBERG, J.S., KONERMANN, S., TREVINO, A., SCOTT, D.A., INOUE, A., MATOBA, S., ZHANG, Y. and ZHANG, F., 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*, **154**(6), pp. 1380-1389.
- ROUET, P., SMIH, F. and JASIN, M., 1994. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, **91**(13), pp. 6064-6068.
- SANDER, J.D. and JOUNG, J.K., 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology*, **32**(4), pp. 347-355.
- SCHUSTER, A. and SCHMOLL, M., 2010. *Biology and biotechnology of Trichoderma*. 87 edn. Applied Microbiology and Biotechnology.
- SEGAL, D.J., DREIER, B., BEERLI, R.R. and BARBAS, C.F., 3rd, 1999. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proceedings of the National Academy of Sciences of the United States of America*, **96**(6), pp. 2758-2763.
- SEIDL, V., GAMAUF, C., DRUZHININA, I.S., SEIBOTH, B., HARTL, L. and KUBICEK, C.P., 2008. The *Hypocrea jecorina* (*Trichoderma reesei*) hypercellulolytic mutant RUT C30 lacks a 85 kb (29 gene-encoding) region of the wild-type genome. *BMC genomics*, **9**, pp. 327-2164-9-327.

SHEN, B., ZHANG, W., ZHANG, J., ZHOU, J., WANG, J., CHEN, L., WANG, L., HODGKINS, A., IYER, V., HUANG, X. and SKARNES, W.C., 2014. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Meth*, **11**(4), pp. 399-402.

STERNBERG, S.H., REDDING, S., JINEK, M., GREENE, E.C. and DOUDNA, J.A., 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, **507**(7490), pp. 62-67.

TEUGJAS, H. and VÄLJAMÄE, P., 2013. Product inhibition of cellulases studied with ¹⁴C-labeled cellulose substrates. *Biotechnology for Biofuels*, **6**(1), pp. 1-14.

VENTER, J.C., ADAMS, M.D., MYERS, E.W., LI, P.W., MURAL, R.J., SUTTON, G.G., SMITH, H.O., YANDELL, M., EVANS, C.A., HOLT, R.A., GOCAYNE, J.D., AMANATIDES, P., BALLEW, R.M., HUSON, D.H., WORTMAN, J.R., ZHANG, Q., KODIRA, C.D., ZHENG, X.H., CHEN, L., SKUPSKI, M., SUBRAMANIAN, G., THOMAS, P.D., ZHANG, J., GABOR MIKLOS, G.L., NELSON, C., BRODER, S., CLARK, A.G., NADEAU, J., MCKUSICK, V.A., ZINDER, N., LEVINE, A.J., ROBERTS, R.J., SIMON, M., SLAYMAN, C., HUNKAPILLER, M., BOLANOS, R., DELCHER, A., DEW, I., FASULO, D., FLANIGAN, M., FLOREA, L., HALPERN, A., HANNENHALLI, S., KRAVITZ, S., LEVY, S., MOBARRY, C., REINERT, K., REMINGTON, K., ABU-THREIDEH, J., BEASLEY, E., BIDDICK, K., BONAZZI, V., BRANDON, R., CARGILL, M., CHANDRAMOULISWARAN, I., CHARLAB, R., CHATURVEDI, K., DENG, Z., DI FRANCESCO, V., DUNN, P., EILBECK, K., EVANGELISTA, C., GABRIELIAN, A.E., GAN, W., GE, W., GONG, F., GU, Z., GUAN, P., HEIMAN, T.J., HIGGINS, M.E., JI, R.R., KE, Z., KETCHUM, K.A., LAI, Z., LEI, Y., LI, Z., LI, J., LIANG, Y., LIN, X., LU, F., MERKULOV, G.V., MILSHINA, N., MOORE, H.M., NAIK, A.K., NARAYAN, V.A., NEELAM, B., NUSSKERN, D., RUSCH, D.B., SALZBERG, S., SHAO, W., SHUE, B., SUN, J., WANG, Z., WANG, A., WANG, X., WANG, J., WEI, M., WIDES, R., XIAO, C., YAN, C., YAO, A., YE, J., ZHAN, M., ZHANG, W., ZHANG, H., ZHAO, Q., ZHENG, L., ZHONG, F., ZHONG, W., ZHU, S., ZHAO, S., GILBERT, D., BAUMHUETER, S., SPIER, G., CARTER, C., CRAVCHIK, A., WOODAGE, T., ALI, F., AN, H., AWE, A., BALDWIN, D., BADEN, H., BARNSTEAD, M., BARROW, I., BEESON, K., BUSAM, D., CARVER, A., CENTER, A., CHENG, M.L., CURRY, L., DANAHER, S., DAVENPORT, L., DESILETS, R., DIETZ, S., DODSON, K., DOUP, L., FERRIERA, S., GARG, N., GLUECKSMANN, A., HART, B., HAYNES, J., HAYNES, C., HEINER, C., HLADUN, S., HOSTIN, D., HOUCK, J., HOWLAND, T., IBEGWAM, C., JOHNSON, J., KALUSH, F., KLINE, L., KODURU, S., LOVE, A., MANN, F., MAY, D., MCCAWLEY, S., MCINTOSH, T., MCMULLEN, I., MOY, M., MOY, L., MURPHY, B., NELSON, K., PFANNKOCH, C., PRATTS, E., PURI, V., QURESHI, H., REARDON, M., RODRIGUEZ, R., ROGERS, Y.H., ROMBLAD, D., RUHFEL, B., SCOTT, R., SITTER, C., SMALLWOOD, M., STEWART, E., STRONG, R., SUH, E., THOMAS, R., TINT, N.N., TSE, S., VEC, C., WANG, G., WETTER, J., WILLIAMS, S., WILLIAMS, M., WINDSOR, S., WINN-DEEN, E., WOLFE, K., ZAVERI, J., ZAVERI, K., ABRIL, J.F., GUIGO, R., CAMPBELL, M.J., SJOLANDER, K.V., KARLAK, B., KEJARIWAL, A., MI, H., LAZAREVA, B., HATTON, T., NARECHANIA, A., DIEMER, K., MURUGANUJAN, A., GUO, N., SATO, S., BAFNA, V., ISTRAIL, S., LIPPERT, R., SCHWARTZ, R., WALENZ, B., YOOSEPH, S., ALLEN, D., BASU, A., BAXENDALE, J., BLICK, L., CAMINHA, M., CARNES-STINE, J., CAULK, P., CHIANG, Y.H., COYNE, M., DAHLKE, C., MAYS, A., DOMBROSKI, M., DONNELLY, M., ELY, D., ESPARHAM, S., FOSLER, C., GIRE, H., GLANOWSKI, S., GLASSER, K., GLODEK, A., GOROKHOV, M., GRAHAM, K., GROPMAN, B., HARRIS, M., HEIL, J., HENDERSON, S., HOOVER, J., JENNINGS, D., JORDAN, C., JORDAN, J., KASHA, J., KAGAN, L., KRAFT, C., LEVITSKY, A., LEWIS, M., LIU, X., LOPEZ, J., MA, D., MAJOROS, W., MCDANIEL, J., MURPHY, S., NEWMAN, M., NGUYEN, T., NGUYEN, N., NODELL, M., PAN, S., PECK, J., PETERSON, M., ROWE, W.,

- SANDERS, R., SCOTT, J., SIMPSON, M., SMITH, T., SPRAGUE, A., STOCKWELL, T., TURNER, R., VENTER, E., WANG, M., WEN, M., WU, D., WU, M., XIA, A., ZANDIEH, A. and ZHU, X., 2001. The sequence of the human genome. *Science (New York, N.Y.)*, **291**(5507), pp. 1304-1351.
- VITIKAINEN, M., ARVAS, M., PAKULA, T., OJA, M., PENTTILÄ, M. and SALOHEIMO, M., 2010. Array comparative genomic hybridization analysis of *Trichoderma reesei* strains with enhanced cellulase production properties. *BMC Genomics*, **11**(1), pp. 1-16.
- VOJTA, A., DOBRINIĆ, P., TADIĆ, V., BOČKOR, L., KORAĆ, P., JULG, B., KLASIĆ, M. and ZOLDOŠ, V., 2016. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic acids research*, .
- WANG, H., YANG, H., SHIVALILA, C.S., DAWLATY, M.M., CHENG, A.W., ZHANG, F. and JAENISCH, R., 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, **153**(4), pp. 910-918.
- WEI, C., LIU, J., YU, Z., ZHANG, B., GAO, G. and JIAO, R., 2013. TALEN or Cas9 – Rapid, Efficient and Specific Choices for Genome Modifications. *Journal of Genetics and Genomics*, **40**(6), pp. 281-289.
- WENINGER, A., HATZL, A., SCHMID, C., VOGL, T. and GLIEDER, A., 2016. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. *Journal of Biotechnology*, , pp. -.
- WU, X., SCOTT, D.A., KRIZ, A.J., CHIU, A.C., HSU, P.D., DADON, D.B., CHENG, A.W., TREVINO, A.E., KONERMANN, S., CHEN, S., JAENISCH, R., ZHANG, F. and SHARP, P.A., 2014. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotech*, **32**(7), pp. 670-676.
- XU, J., ZHAO, G., KOU, Y., ZHANG, W., ZHOU, Q., CHEN, G. and LIU, W., 2014. *Intracellular β -glucosidases CEL1a and CEL1b are essential for cellulase induction on lactose in Trichoderma reesei*. *Eukaryotic Cell*.
- YANG, B. and WYMAN, C.E., 2008. Pretreatment: the key to unlocking low-cost cellulosic ethanol. *Biofuels, Bioproducts and Biorefining*, **2**(1), pp. 26-40.
- YANG, L., GRISHIN, D., WANG, G., AACH, J., ZHANG, C., CHARI, R., HOMSY, J., CAI, X., ZHAO, Y., FAN, J., SEIDMAN, C., SEIDMAN, J., PU, W. and CHURCH, G., 2014. Targeted and genome-wide sequencing reveal single nucleotide variations impacting specificity of Cas9 in human stem cells. *Nat Commun*, **5**.
- ZHANG, W., KOU, Y., XU, J., CAO, Y., ZHAO, G., SHAO, J., WANG, H., WANG, Z., BAO, X., CHEN, G. and LIU, W., 2013. Two Major Facilitator Superfamily Sugar Transporters from *Trichoderma reesei* and Their Roles in Induction of Cellulase Biosynthesis. *Journal of Biological Chemistry*, **288**(46), pp. 32861-32872.
- ZURIS, J.A., THOMPSON, D.B., SHU, Y., GUILINGER, J.P., BESSEN, J.L., HU, J.H., MAEDER, M.L., JOUNG, J.K., CHEN, Z.Y. and LIU, D.R., 2015. Cationic lipid-mediated delivery of proteins

enables efficient protein-based genome editing in vitro and in vivo. *Nature biotechnology*, **33**(1), pp. 73-80.